

# Control and mixing for 2D Navier–Stokes equations with space-time localised noise

Armen Shirikyan\*

October 5, 2011

## Abstract

We consider randomly forced 2D Navier–Stokes equations in a bounded domain with smooth boundary. It is assumed that the random perturbation is non-degenerate, and its law is periodic in time and has a support localised with respect to space and time. Concerning the unperturbed problem, we assume that it is approximately controllable in infinite time by an external force whose support is included in that of the random force. Under these hypotheses, we prove that the Markov process generated by the restriction of solutions to the instants of time proportional to the period possesses a unique stationary distribution, which is exponentially mixing. The proof is based on a coupling argument, a local controllability property of the Navier–Stokes system, an estimate for the total variation distance between a measure and its image under a smooth mapping, and some classical results from the theory of optimal transport.

**AMS subject classifications:** 35Q30, 60H15, 60J05, 93B05, 93C20

**Keywords:** Navier–Stokes equations, stationary measures, exponential mixing

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Main result and scheme of its proof</b>	<b>6</b>
2.1	Exponential mixing . . . . .	6
2.2	General criterion for mixing and application . . . . .	9
<b>3</b>	<b>Control problem</b>	<b>11</b>
3.1	Squeezing . . . . .	11
3.2	Minimisation problem . . . . .	14
3.3	Proof of Theorem 3.1 . . . . .	16

---

\*Department of Mathematics, University of Cergy–Pontoise, CNRS UMR 8088, 2 avenue Adolphe Chauvin, 95302 Cergy–Pontoise, France; E-mail: Armen.Shirikyan@u-cergy.fr

<b>4</b>	<b>Proof of Theorem 2.1</b>	<b>17</b>
4.1	Reduction to a compact phase space . . . . .	18
4.2	Proof of Proposition 2.5 . . . . .	18
4.3	Recurrence . . . . .	19
4.4	Exponential squeezing . . . . .	21
<b>5</b>	<b>Appendix</b>	<b>22</b>
5.1	Optimal coupling . . . . .	22
5.2	Truncated observability inequality . . . . .	24
5.3	Minima of second-order polynomials on Hilbert spaces . . . . .	26
5.4	Image of measures under finite-dimensional transformations . . . . .	26
	<b>Bibliography</b>	<b>28</b>

## 1 Introduction

The main results of this paper can be summarised as follows: first, suitable controllability properties of a non-linear PDE imply the uniqueness and exponential mixing for the associated stochastic dynamics and, second, these properties are satisfied for 2D Navier–Stokes equations with space-time localised noise. To be precise, let us consider from the very beginning the 2D Navier–Stokes system in a bounded domain  $D \subset \mathbb{R}^2$  with smooth boundary  $\partial D$ :

$$\begin{aligned} \dot{u} + \langle u, \nabla \rangle u - \nu \Delta u + \nabla p &= f(t, x), \quad \operatorname{div} u = 0, \quad x \in D, \\ u|_{\partial D} &= 0. \end{aligned} \quad (1.1)$$

Here  $u = (u_1, u_2)$  and  $p$  are unknown velocity and pressure of the fluid,  $\nu > 0$  is the viscosity, and  $f$  is an external force. Let us assume that  $f$  is represented as the sum of two functions  $h$  and  $\eta$ , the first of which is a given function that is  $H^1$  smooth in space and time and has a locally bounded norm, while the second is either a control or a random force:

$$f(t, x) = h(t, x) + \eta(t, x). \quad (1.3)$$

In both cases, we assume that  $\eta$  is sufficiently smooth and bounded, and its restriction to any cylinder of the form  $J_k \times D$  with  $J_k = [k-1, k]$  is localised in both space and time (see below for a more precise description of this hypothesis). Let us denote by  $\mathbf{n}$  the outward unit normal to the boundary  $\partial D$  and introduce the space

$$H = \{u \in L^2(D, \mathbb{R}^2) : \operatorname{div} u = 0 \text{ in } D, \langle u, \mathbf{n} \rangle = 0 \text{ on } \partial D\}, \quad (1.4)$$

which will be endowed with the usual  $L^2$  norm  $\|\cdot\|$ . It is well known that for any  $u_0 \in H$  problem (1.1), (1.2) supplemented with the initial condition

$$u(0, x) = u_0(x) \quad (1.5)$$

has a unique solution  $u = u(t; u_0, f)$ , which is a continuous function of time valued in  $H$ .

To state our first result, we introduce a localisation hypothesis on the control  $\eta$ . Let  $Q \subset J_1 \times D$  be an open set and let  $Q_k = \{(t, x) : (t - k + 1, x) \in Q\}$ .

*Localisation.* For any integer  $k \geq 1$ , the restriction of  $\eta$  to the cylinder  $J_k \times D$  is supported by  $Q_k$ .

The theorem below on the stabilisation to a non-stationary solution for Navier–Stokes equations is an immediate consequence of a contraction property established in Theorem 3.1.

**Theorem A.** *For any  $\rho > 0$  and  $\alpha > 0$  there is a finite-dimensional subspace  $E \subset H_0^1(Q, \mathbb{R}^2)$  and positive constants  $C$  and  $d$  such that the following assertions hold for any functions  $\hat{u}_0 \in H$  and  $h \in H_{\text{loc}}^1(\mathbb{R}_+ \times D, \mathbb{R}^2)$  satisfying the inequalities*

$$\|\hat{u}_0\| \leq \rho, \quad \|h\|_{H^1(J_k \times D)} \leq \rho \quad \text{for all } k \geq 1.$$

- (i) *For any  $u_0 \in H$  satisfying the condition  $\|u_0 - \hat{u}_0\| \leq d$  there is a control  $\eta$  such that the restriction of  $\eta(t + k - 1, x)$  to  $J_1 \times D$  belongs to  $E$  for any integer  $k \geq 1$  and*

$$\|u(t; u_0, h + \eta) - u(t; \hat{u}_0, h)\| \leq Ce^{-\alpha t} \|u_0 - \hat{u}_0\|, \quad t \geq 0. \quad (1.6)$$

- (ii) *The mapping  $(\hat{u}_0, u_0, h) \mapsto \eta$  is Lipschitz continuous for appropriate norms. Moreover,*

$$\|\eta\|_{L^2(Q_k, \mathbb{R}^2)} \leq Ce^{-\alpha k} \|u_0 - \hat{u}_0\|, \quad k \geq 1. \quad (1.7)$$

Let us note that a similar result was proved earlier in [BRS11] for the 3D Navier–Stokes system (see also [Fur04, Bar03, BT04] for some results on stabilisation to a stationary solution). Here we establish some additional properties for the control. Namely, we show that it is finite-dimensional in both space and time and can be chosen to be a smooth function with respect to the force and the initial state corresponding to the reference solution.

We now turn to the stochastic problem. Informally speaking, we assume that the deterministic force  $h$  is 1-periodic in time, and the stochastic force  $\eta$  satisfies the following three conditions. Let us denote by  $\eta_k$  the restriction of  $\eta(t + k - 1, x)$  to the domain  $J_1 \times D$ .

*Independence.* The functions  $\eta_k$  form a sequence of i.i.d. random variables in  $H^1(J_1 \times D, \mathbb{R}^2)$ .

*Non-degeneracy.* There is a compact subset  $\mathcal{K} \subset H_0^1(Q, \mathbb{R}^2)$  such that the support of the law for  $\eta_k$  coincides with  $\mathcal{K}$ , and the law of the projection of  $\eta_k$  to any finite-dimensional subspace of  $L^2(Q, \mathbb{R}^2)$  has a regular density with respect to the Lebesgue measure.

*Approximate controllability.* There is  $\hat{u} \in H$  such that problem (1.1)–(1.3) is approximately controllable to  $\hat{u}$  with a control function  $\tilde{\eta}$  such that for any  $k \geq 1$  the restriction of  $\tilde{\eta}(t+k-1, x)$  to  $J_1 \times D$  belongs to  $\mathcal{K}$ , and the time of control can be chosen the same for the initial functions  $u_0$  from a given bounded subset.

We refer the reader to Section 2.1 for the exact description of the hypotheses imposed on the stochastic system (1.1)–(1.3). Note that the hypothesis on independence implies that the restrictions of solutions to the integer times form a Markov chain in  $H$ . The following theorem describes the long-time asymptotics of this chain.

**Theorem B.** *Under the above hypotheses, the Markov chain associated with problem (1.1)–(1.3) has a unique stationary measure  $\mu$ . Moreover, there are positive constants  $C$  and  $\gamma$  such that, for any 1-Lipschitz function  $F : H \rightarrow \mathbb{R}$  and any  $u_0 \in H$ , we have*

$$\left| \mathbb{E} F(u(k; u_0, h + \eta)) - \int_H F(v) \mu(dv) \right| \leq C(1 + \|u_0\|) e^{-\gamma k}, \quad k \geq 0. \quad (1.8)$$

Let us note that the problem of ergodicity was studied intensively in the last fifteen years. First results in this direction were established in [FM95, KS00, EMS01, BKL02], and we refer the reader to the book [KS11] for further references. Most of the results established so far concern the situation in which the random force is non-degenerate in all determining modes of the problem. In the case when the equation is studied on the torus and the deterministic force is zero, it was proved in [HM06, HM08] that the Navier–Stokes dynamics is exponentially mixing for any  $\nu > 0$ , provided that the noise is white in time and has a few non-zero Fourier modes as a function of  $x$  (thus it is finite-dimensional in space and infinite-dimensional in time). Theorem B applies to all  $\nu > 0$ , and the random noise does not act directly on the deterministic modes. However, in contrast to [HM06, HM08], our noise is localised not in the Fourier modes with respect to  $x$ , but in the physical space and in time. Furthermore, our approach does not use the Malliavin calculus and is based on a detailed study of controllability properties for the Navier–Stokes system and a result on the image of measures on a Hilbert space under finite-dimensional transformations. To the best of my knowledge, Theorem B provides a first result on mixing properties for Navier–Stokes equations with a space-time localised noise.

In conclusion, let us mention that Theorems A and B remain valid in the case when the control and the noise act through the boundary of the domain. This situation will be addressed in a subsequent publication.

The paper is organised as follows. In Section 2, we formulate the main result of this paper on exponential mixing for the Navier–Stokes system with space-time localised noise and outline its proof. Section 3 is devoted to studying a control problem associated with the stochastic system in question. The details of proof of the main result are given in Section 4. The appendix gathers some auxiliary results used in the main text.

## Notation

For an open set  $Q$  of a Euclidean space, a closed interval  $J \subset \mathbb{R}$ , and Banach spaces  $X \subset Y$ , we introduce the following function spaces.

$L^p = L^p(Q)$  is the Lebesgue space of measurable scalar or vector functions on  $Q$  whose  $p^{\text{th}}$  power is integrable. We shall sometimes write  $L^p(Q, \mathbb{R}^d)$  to emphasise the range of functions. In the case  $p = 2$ , the corresponding norm will be denoted by  $\|\cdot\|$ .

$H^s = H^s(Q)$  is the Sobolev space of order  $s$  with the usual norm  $\|\cdot\|_s$ . As in the previous case, we use the same notation for spaces of scalar and vector functions.

$H_0^s = H_0^s(Q)$  is the closure in  $H^s$  of the space of infinitely smooth functions with compact support.

$B_X(R)$  stands for the ball in  $X$  of radius  $R$  centred at zero.

$L^p(J, X)$  is the space of Borel-measurable functions  $u : J \rightarrow X$  such that

$$\|u\|_{L^p(J, X)} = \left( \int_J \|u(t)\|_X^p dt \right)^{1/p} < \infty;$$

in the case  $p = \infty$ , this norm should be replaced by  $\|u\|_\infty = \text{ess sup}_J \|u(t)\|_X$ .

$W^{1,p}(J, X)$  is the space of functions  $u \in L^p(J, X)$  whose derivative belongs to  $L^p(J, X)$ . It is endowed with a natural norm.

$W(J, X, Y)$  is the space of functions  $u \in L^2(J, X)$  such that  $\partial_t u \in L^2(J, Y)$ .

$\mathcal{L}(X, Y)$  is the space of continuous linear operators from  $X$  to  $Y$  with the natural norm. In the case  $X = Y$ , we write  $\mathcal{L}(X)$ .

$C_b(X)$  stands for the space of bounded continuous functions  $F : X \rightarrow \mathbb{R}$ ; it is endowed with the norm

$$\|F\|_\infty = \sup_{u \in X} |F(u)|.$$

$L_b(X)$  denotes the space of functions  $F \in C_b(X)$  such that

$$\|F\|_L := \|F\|_\infty + \sup_{u \neq v} \frac{|F(u) - F(v)|}{\|u - v\|_X} < \infty.$$

$\mathcal{P}(X)$  is the set of probability Borel measures on  $X$ . The space  $\mathcal{P}(X)$  is endowed with the topology of weak convergence, which is generated by the dual-Lipschitz metric

$$\|\mu_1 - \mu_2\|_L^* := \sup_{\|F\|_L \leq 1} |(F, \mu_1) - (F, \mu_2)|,$$

where  $(F, \mu)$  stands for the integral of  $F$  over  $X$  with respect to  $\mu$ .

We denote by  $D$  a bounded domain with  $C^2$  boundary  $\partial D$ . For  $T > 0$ , we set  $J_T = [0, T]$  and  $D_T = J_T \times D$ . The following functional spaces arise in the

theory of Navier–Stokes equations:

$$V = H \cap H_0^1(D), \quad \mathcal{X}_T = W(J_T, V, V^*),$$

$$\mathcal{Y}_{\delta, T} = \{u \in \mathcal{X}_T : u|_{(\delta, T)} \in W((\delta, T), V \cap H^3, V)\},$$

where  $\delta \in (0, T)$ ,  $H$  is defined by (1.4), and  $V^*$  denotes the dual space of  $V$  (identified with a quotient space in  $H^{-1}(D, \mathbb{R}^2)$  with the help of the scalar product in  $L^2$ ). These spaces are endowed with natural norms.

## 2 Main result and scheme of its proof

### 2.1 Exponential mixing

Let  $D \subset \mathbb{R}^2$  be a bounded domain with a  $C^2$ -smooth boundary  $\partial D$  and let  $D_1 = J_1 \times D$ . Consider the Navier–Stokes system (1.1) with the Dirichlet boundary condition (1.2) and an external force of the form (1.3). We assume that  $h \in H_{\text{loc}}^1(\mathbb{R}_+ \times D, \mathbb{R}^2)$  is a given function which is 1-periodic in time and  $\eta$  is a stochastic process of the form

$$\eta(t, x) = \sum_{k=1}^{\infty} I_k(t) \eta_k(t - k + 1, x), \quad t \geq 0, \quad (2.1)$$

where  $I_k$  is the indicator function of the interval  $(k-1, k)$  and  $\{\eta_k\}$  is a sequence of i.i.d. random variables in  $L^2(D_1, \mathbb{R}^2)$  that are continued by zero for  $t \notin J_1$ . It is well known that the Cauchy problem for (1.1)–(1.3) is globally well posed. Namely, for any initial function  $u_0 \in H$  there is a unique random process  $(u, p)$  whose almost every trajectory satisfies the inclusions

$$u \in C(\mathbb{R}_+, H) \cap L_{\text{loc}}^2(\mathbb{R}_+, V), \quad p \in L_{\text{loc}}^1(\mathbb{R}_+, L^2),$$

Equations (1.1), and the initial condition

$$u(0, x) = u_0(x), \quad x \in D. \quad (2.2)$$

In what follows, we shall drop the  $p$  component of solutions and write simply  $u(t)$ . Let us denote by  $S : H \times L^2(D_1, \mathbb{R}^2) \rightarrow H$  the operator that takes a pair of functions  $(u_0, f)$  to  $u(1)$ , where  $u(t)$  is the solution of (1.1), (1.2), (2.2). Well-known properties of 2D Navier–Stokes equations imply that  $S$  is a continuous mapping. Moreover, the range of  $S$  is contained in  $V$ , and the mapping  $S : H \times L^2(D_1, \mathbb{R}^2) \rightarrow V$  is uniformly Lipschitz continuous on bounded subsets; e.g., see Chapter III in [Tem79].

Let us consider a solution  $u(t)$  of (1.1)–(1.3) and denote  $u_k = u(k)$ . What has been said implies that

$$u_k = S(u_{k-1}, h + \eta_k), \quad k \geq 1. \quad (2.3)$$

Since  $\eta_k$  are i.i.d. random variables in  $L^2(D_1, \mathbb{R}^2)$ , Eq. (2.3) defines a homogeneous family of Markov chains in  $H$ , which is denoted by  $(u_k, \mathbb{P}_u)$ ,  $u \in H$ . Let

$P_k(u, \Gamma)$  be the transition function for the family  $(u_k, \mathbb{P}_u)$  and let  $\mathfrak{P}_k : C_b(H) \rightarrow C_b(H)$  and  $\mathfrak{P}_k^* : \mathcal{P}(H) \rightarrow \mathcal{P}(H)$  be the corresponding Markov semigroups.

Let us fix an open set  $Q \subset D_1$  and denote by  $\{\varphi_j\} \subset H^1(Q, \mathbb{R}^2)$  an orthonormal basis in  $L^2(Q, \mathbb{R}^2)$ . Let  $\chi \in C_0^\infty(Q)$  be a non-zero function and let  $\psi_j = \chi \varphi_j$ . In what follows, we shall assume that  $\{\psi_j\}$  are linearly independent<sup>1</sup>, and the function  $h$  and random process  $\eta$  satisfy hypotheses (H1) and (H2) formulated below. Note that if  $\{\varphi_j\}$  is a complete set of eigenfunctions of the Dirichlet Laplacian in  $Q$ , then the functions  $\psi_j$  are linearly independent for any choice of  $\chi$ . This is an immediate consequence of the unique continuation property of solutions for elliptic equations; see Theorem 8.9.1 in [Hör63].

**(H1) Structure of the noise.** The random variables  $\eta_k$  can be represented in the form

$$\eta_k(t, x) = \sum_{j=1}^{\infty} b_j \xi_{jk} \psi_j(t, x), \quad (2.4)$$

where  $\xi_{jk}$  are independent scalar random variables such that  $|\xi_{jk}| \leq 1$  with probability 1, and  $\{b_j\} \subset \mathbb{R}$  is a non-negative sequence such that

$$B := \sum_{j=1}^{\infty} b_j \|\psi_j\|_1 < \infty. \quad (2.5)$$

Moreover, the law of  $\xi_{jk}$  possesses a  $C^1$ -smooth density  $\rho_j$  with respect to the Lebesgue measure on the real line.

Let us denote by  $\mathcal{K} \subset L^2(Q, \mathbb{R}^2)$  the support of the law of  $\eta_k$ . The hypotheses imposed on  $\eta_k$  imply that  $\mathcal{K}$  is a compact subset in  $H_0^1(Q, \mathbb{R}^2)$ . Continuing the elements of  $\mathcal{K}$  by zero outside  $Q$ , we may regard  $\mathcal{K}$  as a compact subset of  $H_0^1(D_1, \mathbb{R}^2)$ .

**(H2) Approximate controllability.** There is  $\hat{u} \in H$  such that for any positive constants  $R$  and  $\varepsilon$  one can find an integer  $l \geq 1$  with the following property: given  $v \in B_H(R)$ , one can find  $\zeta_1, \dots, \zeta_l \in \mathcal{K}$  such that

$$\|S_l(v, \zeta_1, \dots, \zeta_l) - \hat{u}\| \leq \varepsilon, \quad (2.6)$$

where  $S_l(v, \zeta_1, \dots, \zeta_l)$  stands for the vector  $u_l$  defined by (2.3) with  $\eta_k = \zeta_k$  and  $u_0 = v$ .

The following theorem, which is the main result of this paper, establishes the uniqueness and exponential mixing of a stationary distribution for the Markov family generated by (2.3).

---

<sup>1</sup>The assumption on the linear independence of  $\{\psi_j\}$  is not really needed, and Theorem 2.1 below remains true without it. We make, however, this assumption to simplify the proof of Proposition 2.5.

**Theorem 2.1.** Assume that  $h \in H_{\text{loc}}^1(\mathbb{R}_+ \times D, \mathbb{R}^2)$  is 1-periodic in time, and Conditions (H1) and (H2) are satisfied. In this case, there is an integer  $N \geq 1$ , depending on  $\|h\|_{H^1(D_1)}$ ,  $B$ , and  $\nu$ , such that if

$$b_j \neq 0 \quad \text{for } j = 1, \dots, N, \quad (2.7)$$

then the following assertions hold.

**Existence and uniqueness.** The Markov family  $(u_k, \mathbb{P}_u)$  has a unique stationary distribution  $\mu \in \mathcal{P}(H)$ .

**Exponential mixing.** There are positive constants  $C$  and  $\gamma$  such that

$$\|P_k(u, \cdot) - \mu\|_L^* \leq C(1 + \|u\|)e^{-\gamma k} \quad \text{for all } u \in H, k \geq 0. \quad (2.8)$$

Note that condition (2.7) expresses the space-time non-degeneracy of the noise. Thus, the property of exponential mixing holds true even for noises whose space-time dimension is finite.

The general scheme of the proof of Theorem 2.1 is outlined in Section 2.2, and the details are given in Sections 3 and 4. Here we consider two examples for which Condition (H2) is fulfilled.

*Example 2.2.* We claim that there is  $\delta > 0$  such that if the 1-periodic function  $h$  satisfies the inequality  $\|h\|_{L^2(D_1)} \leq \delta$ , then Condition (H2) is fulfilled, provided that  $\mathcal{K}$  contains the zero element.

Indeed, if  $\|h\|_{L^2(D_1)}$  is sufficiently small, then problem (1.1), (1.2) with  $f = h$  has a unique solution  $\tilde{u}(t, x)$  defined throughout the real line and 1-periodic in time. To see this, it suffices to take a sequence  $\{u_n\}$  of solutions for the problem in question such that  $u_n(-n) = 0$  and to prove that it converges as  $n \rightarrow \infty$  in the space  $W([-N, N], V, V^*)$  for any  $N > 0$ . The limiting function  $\tilde{u}$  is the required 1-periodic solution. Using standard estimates for Navier–Stokes equations, it is easy to prove that

$$\|\tilde{u}\|_{W(J_1, V, V^*)} \leq c(\delta) \rightarrow 0 \quad \text{as } \delta \rightarrow 0.$$

This implies that  $\tilde{u}$  is globally exponentially stable as  $t \rightarrow +\infty$ . Therefore, for any positive constants  $R$  and  $\varepsilon$  one can find an integer  $l \geq 1$  such that (2.6) holds with  $\zeta_1 = \dots = \zeta_l = 0$  and  $\hat{u} = \tilde{u}(0)$ . Since  $\mathcal{K}$  contains the zero element, we see that Condition (H2) is satisfied.

*Example 2.3.* Suppose that  $h$  is represented in the form  $h(t, x) = h_0(\lambda t, x)$ , where  $h_0$  is a continuous 1-periodic function of time with range in  $V$  such that  $\int_0^1 h_0(s) ds = 0$ . We claim that, for sufficiently large integers  $\lambda > 0$ , Condition (H2) is satisfied, provided that  $\mathcal{K}$  contains the zero element.

Indeed, let us represent a solution of Eq. (1.1) with  $f = h$  in the form  $u = w + g$ , where  $g(t) = \int_0^t h(s) ds$ . Then the function  $w$  must satisfy the equations

$$\dot{w} + \langle w + g, \nabla \rangle (w + g) - \nu \Delta w + \nabla p = -\Delta g, \quad \text{div } w = 0. \quad (2.9)$$

The condition imposed on  $h$  imply that  $g$  is a  $V$ -valued 1-periodic function such that

$$\sup_{t \in \mathbb{R}} \|g(t)\|_V \rightarrow 0 \quad \text{as } \lambda \rightarrow \infty.$$

Combining this with an argument similar to that used in Example 2.2, one can prove that Eq. (2.9) has a unique 1-periodic solution  $\tilde{w}$ , which is globally exponentially stable as  $t \rightarrow \infty$ . It follows that  $\tilde{u} = \tilde{w} + g$  is a globally exponentially stable solution for problem (1.1), (1.2) with  $f = h$ . As in Example 2.2, we conclude that the function  $\hat{u} = \tilde{u}(0)$  satisfies the required property.

## 2.2 General criterion for mixing and application

In this section, we outline the proof of Theorem 2.1, which is based on two key ingredients: a coupling approach developed in [KS01, Mat02, KPS02, MY02, Hai02, Shi04] in the context of stochastic PDE's and a property of stabilisation to a non-stationary solution of Navier–Stokes equations [BRS11]. We first recall an abstract result established in [Shi08].

Let  $X$  be a compact metric space with a metric  $d_X$  and let  $(u_k, \mathbb{P}_u)$ ,  $u \in X$ , be a family of Markov chains in  $X$ . We denote by  $P_k(u, \Gamma)$  its transition function and by  $\mathfrak{P}_k$  and  $\mathfrak{P}_k^*$  the corresponding Markov semigroups. Let  $(\mathbf{u}_k, \mathbb{P}_{\mathbf{u}})$  be another family of Markov chains in the extended phase space  $\mathbf{X} = X \times X$  such that

$$\Pi_* \mathbf{P}_k(\mathbf{u}, \cdot) = P_k(u, \cdot), \quad \Pi'_* \mathbf{P}_k(\mathbf{u}, \cdot) = P_k(u', \cdot) \quad \text{for } \mathbf{u} = (u, u') \in \mathbf{X}, \quad k \geq 0, \quad (2.10)$$

where  $\mathbf{P}_k(\mathbf{u}, \Gamma)$  denotes the transition function for  $(\mathbf{u}_k, \mathbb{P}_{\mathbf{u}})$  and  $\Pi, \Pi' : \mathbf{X} \rightarrow X$  stand for the natural projections to the components of a vector  $\mathbf{u} = (u, u')$ . In other words, relations (2.10) mean that, for any integer  $k \geq 1$ , the random variable  $\mathbf{u}_k$  considered under the law  $\mathbb{P}_{\mathbf{u}}$  with  $\mathbf{u} = (u, u')$  is a coupling for the pair of measures  $(P_k(u, \cdot), P_k(u', \cdot))$ . We shall say that  $(\mathbf{u}_k, \mathbb{P}_{\mathbf{u}})$  satisfies the *mixing hypothesis* if there is a closed subset  $\mathbf{B} \subset \mathbf{X}$  and positive constants  $C$  and  $\beta$  such that the following properties hold.

**Recurrence.** Let  $\tau(\mathbf{B})$  be the first hitting time of the set  $\mathbf{B}$ :

$$\tau(\mathbf{B}) = \min\{k \geq 0 : \mathbf{u}_k \in \mathbf{B}\}.$$

Then  $\tau(\mathbf{B})$  is  $\mathbb{P}_{\mathbf{u}}$ -almost surely finite for any  $\mathbf{u} \in \mathbf{X}$ , and there are positive constants  $C_1$  and  $\delta_1$  such that

$$\mathbb{E}_{\mathbf{u}} \exp(\delta_1 \tau(\mathbf{B})) \leq C_1 \quad \text{for } \mathbf{u} \in \mathbf{X}. \quad (2.11)$$

**Exponential squeezing.** Let us set

$$\sigma = \min\{k \geq 0 : d_X(u_k, u'_k) > C e^{-\beta k}\}. \quad (2.12)$$

Then there are positive constants  $C_2$ ,  $\delta_2$ , and  $\delta_3$  such that, for any  $\mathbf{u} \in \mathbf{B}$ , we have

$$\mathbb{P}_{\mathbf{u}}\{\sigma = \infty\} \geq \delta_3, \quad (2.13)$$

$$\mathbb{E}_{\mathbf{u}}(I_{\{\sigma < \infty\}} \exp(\delta_2 \sigma)) \leq C_2. \quad (2.14)$$

The following proposition is a particular case of a more general result established<sup>2</sup> in [Shi08] (see Theorem 2.3).

**Proposition 2.4.** *Let  $(u_k, \mathbb{P}_u)$  be a family of Markov chains for which there exists another Markov family  $(\mathbf{u}_k, \mathbb{P}_{\mathbf{u}})$  in the extended space  $\mathbf{X}$  that satisfies relation (2.10) and the mixing hypothesis. Then  $(u_k, \mathbb{P}_u)$  has a unique stationary distribution  $\mu \in \mathcal{P}(X)$ , and there are positive constants  $C$  and  $\gamma$  such that*

$$\|P_k(u, \cdot) - \mu\|_L^* \leq C e^{-\gamma k} \quad \text{for } u \in X, k \geq 0. \quad (2.15)$$

To prove Theorem 2.1, we first observe that the Markov family  $(u_k, \mathbb{P}_u)$  possesses a compact absorbing invariant set  $X \subset H$ , and it suffices to study its restriction to  $X$ , for which we retain the same notation. We shall prove that  $(u_k, \mathbb{P}_u)$  satisfies the hypotheses of Proposition 2.4. A crucial point of our construction is the following result, which says, roughly speaking, that if two points  $u, u' \in X$  are sufficiently close, then the pair  $(P_1(u, \cdot), P_1(u', \cdot))$  admits a coupling whose components are close with high probability; cf. Lemma 3.3 in [KS01].

**Proposition 2.5.** *Under the hypotheses of Theorem 2.1, there exists a constant  $d > 0$  such that for any points  $u, u' \in X$  satisfying the inequality  $\|u - u'\| \leq d$  the pair  $(P_1(u, \cdot), P_1(u', \cdot))$  admits a coupling  $(V(u, u'), V'(u, u'))$  such that*

$$\mathbb{P}\{\|V(u, u') - V'(u, u')\| > \tfrac{1}{2}\|u - u'\|\} \leq C\|u - u'\|, \quad (2.16)$$

where  $C > 0$  is a constant not depending on  $u, u' \in X$ .

The proof of this proposition is based on a controllability property for the Navier–Stokes system and application of a concept of optimal coupling; see Sections 3 and 5.1. We now define a *coupling operator*  $\mathcal{R} = (\mathcal{R}, \mathcal{R}')$  by the relation

$$\mathcal{R}(u, u'; \omega) = \begin{cases} (V(u, u'), V'(u, u')) & \text{for } \|u - u'\| \leq d, \\ (S(u, \zeta), S(u', \zeta')) & \text{for } \|u - u'\| > d, \end{cases} \quad (2.17)$$

where  $\zeta$  and  $\zeta'$  are independent random variables whose law coincides with that of  $\eta_1$ . Without loss of generality, we can assume that  $\zeta$  and  $\zeta'$  are defined on the same probability space as  $V$  and  $V'$ . The required Markov family  $(\mathbf{u}_k, \mathbb{P}_{\mathbf{u}})$  is constructed by iterations of  $\mathcal{R}$ . Namely, let  $(\Omega_k, \mathcal{F}_k, \mathbb{P}_k)$ ,  $k \geq 1$ , be countably

<sup>2</sup>In [Shi08], the proof is carried out in the particular case when  $\mathbf{B} = B \times B$ ; however, the same argument applies in the general situation.

many copies of the probability space on which  $\mathcal{R}$  is defined and let  $(\Omega, \mathcal{F}, \mathbb{P})$  be the direct product of these spaces. We set

$$\mathbf{u}_0 = (u, u'), \quad \mathbf{u}_k = \mathcal{R}(\mathbf{u}_{k-1}, \omega_k), \quad k \geq 1. \quad (2.18)$$

The recurrence property will follow from approximate controllability (see Hypothesis (H2) in Section 2.1), while the exponential squeezing will be implied by Proposition 2.5.

### 3 Control problem

#### 3.1 Squeezing

In this section, we consider the controlled Navier–Stokes system (1.1)–(1.3) on the time interval  $J_1 = [0, 1]$ . We assume that the function  $h$  belongs to the space  $H^1(D_1)$  (where  $D_1 = J_1 \times D$ ) and denote by  $\{\psi_j\}$  the sequence of functions entering Hypothesis (H1) of Section 2.1. Extending the functions  $\psi_j$  by zero outside  $Q$ , we may regard them as elements of  $H_0^1(D_1)$ . We denote by  $\mathcal{E}_m$  the vector span of  $\psi_1, \dots, \psi_m$  endowed with the  $L^2$  norm and by  $B_R$  the ball in  $H^1(D_1)$  of radius  $R$  centred at origin.

**Theorem 3.1.** *Under the above hypotheses, for any  $R > 0$  and  $q \in (0, 1)$  there is an integer  $m \geq 1$ , positive constants  $d$  and  $C$ , and a continuous mapping*

$$\Phi : B_R \times B_H(R) \rightarrow \mathcal{L}(H, \mathcal{E}_m), \quad (h, \hat{u}_0) \mapsto \eta,$$

*such that the following properties hold.*

**Contraction.** *For any functions  $h \in B_R$  and  $\hat{u}_0, u_0 \in B_H(R)$  satisfying the inequality*

$$\|u_0 - \hat{u}_0\| \leq d, \quad (3.1)$$

*we have*

$$\|S(\hat{u}_0, h) - S(u_0, h + \Phi(h, \hat{u}_0)(u_0 - \hat{u}_0))\| \leq q \|u_0 - \hat{u}_0\|. \quad (3.2)$$

**Regularity.** *The mapping  $\Phi$  is infinitely smooth in the Fréchet sense.*

**Lipschitz continuity.** *The mapping  $\Phi$  is Lipschitz continuous with the constant  $C$ . That is,*

$$\|\Phi(h_1, \hat{u}_1) - \Phi(h_2, \hat{u}_2)\|_{\mathcal{L}} \leq C (\|h_1 - h_2\|_{H^1} + \|\hat{u}_1 - \hat{u}_2\|), \quad (3.3)$$

*where  $\|\cdot\|_{\mathcal{L}}$  stands for the norm in the space  $\mathcal{L}(H, \mathcal{E}_m)$ .*

Taking this result for granted, let us prove Theorem A stated in the Introduction. Let us fix positive constants  $\rho$  and  $\alpha$  and take an initial function

$\hat{u}_0 \in B_H(\rho)$ . The boundedness of the resolving operator for the Navier–Stokes system implies that the corresponding solution  $u$  satisfies the inequality

$$\|u(t)\| \leq R \quad \text{for all } t \geq 0, \quad (3.4)$$

where  $R$  is a constant depending only on  $\rho$ . Let  $q \in (0, 1)$  be such that  $e^{-\alpha} = q$ . Denote by  $d$ ,  $C$ , and  $m$  the parameters constructed in Theorem 3.1. Let us take any  $u_0 \in H$  satisfying the inequality  $\|u_0 - \hat{u}_0\| \leq d$  and construct a control function  $\eta$  consecutively on the intervals  $J_k$ ,  $k \geq 1$ . To this end, we denote by  $h_k$  the restriction of the function  $h(t - k + 1)$  to  $D_1$  and set

$$\eta(t) = \Phi(\hat{u}_0, h_1)(u_0 - \hat{u}_0) \quad \text{for } 0 \leq t \leq 1.$$

Then, by Theorem 3.1, we have

$$\|u(1) - \hat{u}(1)\| \leq q\|u_0 - \hat{u}_0\| = e^{-\alpha}\|u_0 - \hat{u}_0\|,$$

where  $u(t)$  denotes the solution of (1.1)–(1.3), (1.5) on the interval  $J_1$ . Assume we have constructed a control  $\eta$  on the interval  $(0, k)$  with  $k \geq 1$ , and the corresponding solution  $u$  satisfies the inequality

$$\|u(l) - \hat{u}(l)\| \leq e^{-\alpha l}\|u_0 - \hat{u}_0\| \quad (3.5)$$

for  $l = 1, \dots, k$ . In this case, setting

$$\eta(k + t) = \Phi(\hat{u}(k), h_{k+1})(u(k) - \hat{u}(k)) \quad \text{for } 0 \leq t \leq 1, \quad (3.6)$$

we see that (3.5) remains valid for  $l = k + 1$ . Thus, we can construct  $\eta$  on the half-line  $\mathbb{R}_+$ , and the corresponding solution will satisfy (3.5) for all integers  $l \geq 1$ . Combining this with the Lipschitz continuity of the resolving operator of the Navier–Stokes system, we see that (1.6) holds. Inequality (1.7) is a straightforward consequence of (3.6) and (3.5). Finally, it is not difficult to check that if  $\eta^1$  and  $\eta^2$  are two controls corresponding to  $(u_0^i, \hat{u}_0^i, h^i)$ ,  $i = 1, 2$ , then

$$\|\eta_k^1 - \eta_k^2\|_{L^2(D_k)} \leq C_1^k \left( \|u_0^1 - u_0^2\| + \|\hat{u}_0^1 - \hat{u}_0^2\| + \max_{1 \leq l \leq k} \|h_l^1 - h_l^2\|_{H^1(D_l)} \right),$$

where  $C_1 > 0$  depends only on  $\rho$  and  $\alpha$ ,  $\eta_l^i$  stands for the restriction of  $\eta^i$  to  $D_l$ , and  $h_l^i$  are defined in a similar way. This completes the proof of Theorem A.

We now turn to the proof of Theorem 3.1. As was mentioned in the Introduction, a weaker version of this result was established in the paper [BRS11], and its proof repeats essentially the argument used there. However, since the finite-dimensionality in time for the control and the regularity and Lipschitz properties of  $\Phi$  are important for the stochastic part of this work, we present a rather complete proof of Theorem 3.1. We begin with a description of the main steps and give the details in the next two subsections.

*Step 1: Reduction to the linearised problem.* Denote by  $\hat{u}(t, x)$  the solution of (1.1), (1.3) issued from  $\hat{u}_0$  and corresponding to  $\eta \equiv 0$ . In view of the regularising property of the Navier–Stokes system, for any interval  $J = (\delta, 1)$  with  $\delta > 0$  we have

$$\hat{u} \in L^2(J, H^3 \cap V), \quad \partial_t \hat{u} \in L^2(J, V), \quad (3.7)$$

and the corresponding norms are bounded by a constant depending only on  $\delta$  and  $R$ . In particular, the truncated observability inequality holds for the adjoint of the Navier–Stokes system linearised around  $\hat{u}$ ; see Section 5.2.

A solution with a non-zero control  $\eta$  is sought in the form  $u = \hat{u} + v$ . Then  $v$  must be a solution of the problem

$$\dot{v} + \langle v, \nabla \rangle v + \langle \hat{u}, \nabla \rangle v + \langle v, \nabla \rangle \hat{u} - \nu \Delta v + \nabla p = \eta(t, x), \quad \operatorname{div} v = 0, \quad (3.8)$$

$$v|_{\partial D} = 0, \quad v(0) = v_0, \quad (3.9)$$

where  $v_0 = u_0 - \hat{u}_0$ . Together with Eq. (3.8), consider its linearisation around zero:

$$\dot{v} + \langle \hat{u}, \nabla \rangle v + \langle v, \nabla \rangle \hat{u} - \nu \Delta v + \nabla p = \eta(t, x), \quad \operatorname{div} v = 0. \quad (3.10)$$

Suppose that we have constructed  $\eta \in \mathcal{E}_m$  such that the solution  $w(t, x)$  of (3.10), (3.9) satisfies the inequalities

$$\|w(1)\| \leq \frac{q}{2} \|v_0\|, \quad \|w\|_{\mathcal{X}_1} \leq C_1 \|v_0\|. \quad (3.11)$$

A standard perturbative argument shows that if  $\|v_0\|$  is sufficiently small, then the solution of (3.8), (3.9) satisfies the inequality  $\|v(1)\| \leq q \|v_0\|$ , whence it follows that (3.2) holds. Thus, it suffices to construct a continuous linear operator  $\Phi(h, \hat{u}_0) : H \rightarrow \mathcal{E}_m$  such that the solution  $w \in \mathcal{X}_1$  of problem (3.10), (3.9) with  $\eta = \Phi(h, \hat{u}_0)v_0$  satisfies inequalities (3.11).

*Step 2: Application of the Foias–Prodi property.* Let  $\{e_j\}$  be an orthonormal basis in  $H$  formed of the eigenfunctions of the Stokes operator  $L = -\Pi\Delta$ , where  $\Pi$  stands for the Leray projection in  $L^2(D, \mathbb{R}^2)$  (onto the closed subspace  $H$ ), let  $\{\alpha_j\}$  be the corresponding (non-decreasing) sequence of eigenvalues for  $L$ , and let  $\Pi_N$  be the orthogonal projection in  $H$  on the vector space  $H_N$  spanned by  $e_1, \dots, e_N$ . Denote by  $\mathcal{R}^{\hat{u}} : H \times L^2(D_1) \rightarrow \mathcal{X}_1$  a linear operator that takes  $(v_0, \eta)$  to the solution  $w$  of (3.10), (3.9) and by  $\mathcal{R}_t^{\hat{u}}$  its restriction to the time  $t$ .

Suppose that for any integer  $N \geq 1$  and any  $\delta > 0$  we have constructed an integer  $m \geq 1$  and a family of linear operators  $\Phi = \Phi(h, \hat{u}_0) : H \rightarrow \mathcal{E}_m$  which is a Lipschitz function of its arguments and is such that

$$\|\Pi_N \mathcal{R}_1^{\hat{u}}(v_0, \Phi(h, \hat{u}_0)v_0)\| \leq C_2 \delta \|v_0\|, \quad \|\Phi(h, \hat{u}_0)\|_{\mathcal{L}} \leq C_2, \quad (3.12)$$

where  $C_2 > 0$  is a constant not depending on  $N$  and  $\delta$ . In this case, the Poincaré

inequality and the regularising property of (3.10) imply that

$$\begin{aligned}
\|\mathcal{R}_1^{\hat{u}}(v_0, \Phi(h, \hat{u}_0)v_0)\| &= \|(I - \Pi_N)\mathcal{R}_1^{\hat{u}}(v_0, \Phi(h, \hat{u}_0)v_0)\| + C_2\delta \|v_0\| \\
&\leq \alpha_{N+1}^{-1/2} \|\mathcal{R}_1^{\hat{u}}(v_0, \Phi(h, \hat{u}_0)v_0)\|_1 + C_2\delta \|v_0\| \\
&\leq C_3\alpha_{N+1}^{-1/2} (\|v_0\| + \|\Phi(h, \hat{u}_0)v_0\|_{L^2(D_1)}) + C_2\delta \|v_0\| \\
&\leq (C_3(C_2 + 1)\alpha_{N+1}^{-1/2} + C_2\delta) \|v_0\|.
\end{aligned}$$

Choosing  $N$  sufficiently large and  $\delta$  sufficiently small, we obtain the first inequality in (3.11). The second is an immediate consequence of the continuity of  $\mathcal{R}^{\hat{u}}$  and the boundedness of  $\Phi$ .

*Step 3: Minimisation problem.* The construction of  $\Phi$  is based on a study of a minimisation problem for solutions of (3.10) with a cost functional penalising the term  $\|\Pi_N w(1)\|$ . Namely, let us consider the following problem.

**Problem 3.2.** Given a constant  $\delta > 0$ , an integer  $N \geq 1$ , and functions  $v_0 \in H$  and  $\hat{u} \in \mathcal{X}_1$  satisfying (3.7), minimise the functional

$$J(w, \zeta) = \frac{1}{2} \int_0^1 \|\zeta(t)\|^2 dt + \frac{1}{\delta} \|\Pi_N w(1)\|^2$$

over the set of functions  $(w, \zeta) \in \mathcal{X}_1 \times L^2(D_1, \mathbb{R}^2)$  satisfying the equations

$$\dot{w} + \langle \hat{u}, \nabla \rangle w + \langle w, \nabla \rangle \hat{u} - \nu \Delta w + \nabla p = \chi(\mathbf{P}_m \zeta), \quad \operatorname{div} w = 0, \quad w(0) = v_0, \quad (3.13)$$

where  $p = p(t, x)$  is a distribution in  $D_1$  and  $\mathbf{P}_m$  stands for the orthogonal projection in  $L^2(D_1)$  onto  $\mathcal{E}_m$ .

We shall show that Problem 3.2 has a unique solution  $(w, \zeta)$ , which satisfies the inequality

$$\frac{1}{\delta} \|\Pi_N w(1)\|^2 + \|\zeta\|_{L^2(D_1)}^2 \leq C_3 \|v_0\|^2, \quad (3.14)$$

where  $C_3 > 0$  is a constant not depending on  $N$  and  $\delta$ . This will imply the required inequalities (3.12), in which  $\Phi v_0 = \chi(\mathbf{P}_m \zeta)$ . Further analysis shows that the mapping  $\hat{u} \mapsto \zeta$  is smooth from  $\mathcal{X}_1$  to  $L^2(D_1)$  and uniformly Lipschitz continuous on bounded balls. Since  $\hat{u}$  is an analytic function of  $(h, \hat{u}_0)$ , this will complete the proof of the theorem.

### 3.2 Minimisation problem

*Step 1: Existence, uniqueness, and linearity.* Let us prove that Problem 3.2 has a unique optimal solution  $(w, \zeta)$  in the space  $X := \mathcal{X}_1 \times L^2(D_1)$ . Indeed, the function  $J : X \rightarrow \mathbb{R}$  is non-negative and therefore has an infimum on any affine subspace of  $X$ . Denote by  $X_{v_0}$  the affine subspace of  $X$  defined by (3.13) and by  $J_*$  the infimum of  $J$  on  $X_{v_0}$ . Let  $(w_n, \zeta_n) \in X_{v_0}$  be an arbitrary minimising sequence. Then  $\{\zeta_n\}$  is a bounded sequence in  $L^2(D_1)$ , and without loss of generality we can assume that it converges weakly to a limit  $\zeta$ . It follows that

the sequence of solutions  $w_n \in \mathcal{X}_1$  of problem (3.13) with  $\zeta = \zeta_n$  converges in the space  $\mathcal{X}_1$  to a limit  $w$ , which satisfies (3.13). The lower semi-continuity of the norm of a Hilbert space now implies that

$$J(w, \zeta) \leq \liminf_{n \rightarrow \infty} J(w_n, \zeta_n) = J_*.$$

Recalling the definition of  $J_*$ , we conclude that  $J(w, \zeta) = J_*$ .

To prove the uniqueness, note that any affine subspace is a convex set in  $X$ . Combining this property with the strict convexity of the norm of a Hilbert space, we see that if  $(w_i, \zeta_i)$ ,  $i = 1, 2$ , are two optimal solutions, then  $\zeta_1 = \zeta_2$ . Since the solution of problem (3.13) with a given  $\zeta \in L^2(D_1)$  is unique, we conclude that  $w_1 = w_2$ .

Finally, it is a standard fact of the optimisation theory that the unique minimum of a quadratic functional under a linear constraint can be expressed as a linear function of the problem data. In the case under study, the corresponding operator depends on the reference solution  $\hat{u}$ . We shall denote by  $\Psi(\hat{u})$  the linear operator that takes  $v_0$  to  $\zeta$ , where  $(w, \zeta)$  is the optimal solution for Problem 3.2.

*Step 2: Regularity and Lipschitz continuity.* We now prove that the mapping  $\Psi$  regarded as an application from  $H$  to  $\mathcal{L}(H, L^2(D_1))$  is infinitely differentiable and uniformly Lipschitz continuous on balls. Let us denote by  $w_0 = w_0(\hat{u}, v_0) \in \mathcal{X}_1$  the solution of problem (3.13) with  $\zeta \equiv 0$ . The linear constraint of Problem 3.2 is equivalent to the relation

$$w = w_0 + \mathcal{R}^{\hat{u}}(0, \chi(\mathbf{P}_m \zeta)). \quad (3.15)$$

Setting  $A(\hat{u})\zeta = \mathcal{R}_1^{\hat{u}}(0, \chi(\mathbf{P}_m \zeta))$ , we see that  $(w, \zeta) \in X$  is a solution of Problem 3.2 if and only if  $w$  is the global minimum of the function  $F : L^2(D_1) \rightarrow \mathbb{R}$  defined by

$$F(\zeta; \hat{u}, v_0) = \frac{1}{2} \int_0^1 \|\zeta(t)\|^2 dt + \frac{1}{\delta} \|\Pi_N(w_0 + A(\hat{u})\zeta)\|^2.$$

Using standard methods of the theory of 2D Navier–Stokes equations, we can prove that  $A(\hat{u})$  is an analytic function from  $\mathcal{X}_1$  to  $\mathcal{L}(H)$  which is uniformly Lipschitz continuous on bounded sets; e.g., see [Kuk82] or Chapter 1 in [VF88]. It follows that  $F$  satisfies the hypotheses of Proposition 5.5 in which  $U = L^2(D_1)$  and  $Y = \mathcal{X}_1 \times H$ . Thus, the unique minimum  $\zeta$  of  $F$  is a smooth function of  $(\hat{u}, v_0)$  valued in  $L^2(D_1)$ , and it is Lipschitz continuous on bounded subsets. Recalling relation (3.15), we conclude that the unique solution  $(w, \zeta) \in X$  of Problem 3.2 is a smooth function on  $\mathcal{X}_1 \times H$  which is Lipschitz continuous on bounded subsets. Since  $(w, \zeta)$  linearly depends on  $v_0$ , it is straightforward to derive the required properties of  $\Psi$ .

*Step 3: A priori estimate.* From now on, we fix  $R, \delta > 0$  and assume that  $\hat{u} \in B_{\mathcal{Y}_{\delta,1}}(R)$ . We claim that there is an integer  $m \geq 1$  depending on  $R$  and  $N$ , and a constant  $C > 0$  depending only on  $R$  such that, if  $\hat{u} \in B_{\mathcal{Y}_{\delta,1}}(R)$ , then

inequality (3.14) holds. Indeed, note that the constraint given by (3.13) is equivalent to the equations

$$\dot{w} + \nu Lw + B(\hat{u}, w) + B(w, \hat{u}) = \Pi(\chi(\mathbf{P}_m \zeta)), \quad w(0) = v_0, \quad (3.16)$$

where we set  $B(u_1, u_2) = \Pi(\langle u_1, \nabla \rangle u_2)$ . Thus, the pair  $(w, \zeta) \in X$  constructed in Step 1 is the minimiser of  $J$  under constraint (3.14). Applying the Kuhn–Tucker theorem (see Chapter I in [IT79]), we can find functions  $\lambda \in H$  and  $\theta \in L^2(J_1, V)$  such that, for any  $(r, \xi) \in X$ , we have

$$\begin{aligned} \int_0^1 (\zeta, \xi) dt + \frac{2}{\delta} (\Pi_N w(1), r(1)) + (\lambda, r(0)) \\ + \int_0^1 (\theta, \dot{r} + \nu Lr + B(\hat{u}, r) + B(r, \hat{u}) - \chi(\mathbf{P}_m \xi)) dt = 0. \end{aligned}$$

It follows that

$$\dot{\theta} - \nu L\theta - B^*(\hat{u})\theta = 0, \quad (3.17)$$

$$\theta(1) = -\frac{2}{\delta} \Pi_N w(1), \quad \zeta = \mathbf{P}_m(\chi\theta), \quad (3.18)$$

where  $B^*(\hat{u})$  denotes the (formal) adjoint of the operator  $B(\hat{u}, \cdot) + B(\cdot, \hat{u})$  in  $H$ , and Eq. (3.17) holds in the sense of distributions. Multiplying Eq. (3.17) by  $w$  and the first equation in (3.16) by  $\theta$ , adding together the resulting relations, and integrating over  $D_1$ , we derive

$$\frac{2}{\delta} \|\Pi_N w(1)\|^2 + \int_0^1 \|\zeta(t)\|^2 dt = -(\theta(0), v_0), \quad (3.19)$$

where we used relations (3.18) and the initial condition in (3.16). Now note that (3.17) is equivalent to the backward Navier–Stokes equations (5.12). Therefore, by Proposition 5.4, the function  $\theta$  must satisfy the truncated observability inequality (5.13). Combining it with (3.19), we obtain (3.14). We have thus shown that (cf. (3.12))

$$\|\Pi_N \mathcal{R}_1^{\hat{u}}(v_0, \chi \mathbf{P}_m(\Psi(\hat{u})v_0))\| \leq C\delta \|v_0\|, \quad \|\Psi(\hat{u})v_0\|_{L^2(D_1)} \leq C\|v_0\|. \quad (3.20)$$

### 3.3 Proof of Theorem 3.1

Let us define  $\Phi(h, \hat{u}_0)$  as the linear operator taking  $v_0$  to  $\chi(\mathbf{P}_m \Psi(\hat{u})v_0)$ , where  $\hat{u} \in \mathcal{X}_1$  is the solution of problem (1.1), (1.6) with  $f = h$  and  $u_0 = \hat{u}_0$ . Standard regularity results for 2D Navier–Stokes equations imply that  $\hat{u} \in \mathcal{Y}_{\delta,1}$  for any  $\delta \in (0, 1)$  (see Theorem 3.5 in Chapter III of [Tem79]), so that  $\Psi(\hat{u})$  is well defined. Moreover, the norm  $\|\hat{u}\|_{\mathcal{Y}_{\delta,1}}$  is bounded by a constant depending only on  $R$  and  $\delta$ . We claim that  $\Phi(h, \hat{u}_0)$  satisfies the contraction property stated in the theorem. If this assertion is proved, then the regularity and Lipschitz continuity of  $\Psi$  combined with similar properties of the resolving operator for the 2D Navier–Stokes system will imply the remaining assertions on  $\Phi$ .

Inequalities (3.20) imply that the solution  $(w, \zeta)$  of Problem 3.2 satisfies (3.14). Therefore, choosing  $N$  and  $\delta^{-1}$  sufficiently large, we ensure that the function  $w = \mathcal{R}^{\hat{u}}(v_0, \Phi(h, \hat{u}_0)v_0)$  satisfies (3.11). Let us represent a solution of the non-linear problem (3.8), (3.9) with the right-hand side  $\eta = \Phi(h, \hat{u}_0)(u_0 - \hat{u}_0)$  in the form  $v = w + z$ . Then  $z \in \mathcal{X}_1$  must be a solution of the problem

$$\dot{z} + \langle z, \nabla \rangle z + \langle \hat{u} + w, \nabla \rangle z + \langle z, \nabla \rangle (\hat{u} + w) - \nu \Delta z + \nabla p = -\langle w, \nabla \rangle w, \quad (3.21)$$

$$\operatorname{div} z = 0, \quad z|_{\partial D} = 0, \quad z(0) = 0. \quad (3.22)$$

Taking the scalar product of (3.21) with  $2z$  in  $H$  and carrying out some standard transformations, we derive

$$\partial_t \|z\|^2 + 2\nu \|\nabla z\|^2 \leq C_1 (\|\hat{u} + w\|_1 \|z\| + \|w\|_1 \|w\|) \|\nabla z\|.$$

It follows that

$$\partial_t \|z\|^2 + \nu \|\nabla z\|^2 \leq C_2 (\|\hat{u} + w\|_1^2 \|z\|^2 + \|w\|_1^2 \|w\|^2).$$

Applying the Gronwall inequality and using the initial condition in (3.22), we obtain

$$\|z(t)\|^2 \leq \int_0^t \exp\left(C_2 \int_s^t \|\hat{u} + w\|_1^2 dr\right) \|w\|_1^2 \|w\|^2 ds.$$

Recalling that  $\|\hat{u}\|_{\mathcal{X}_1}$  is bounded by a constant depending only on  $R$  and using the second inequality in (3.11), we derive

$$\sup_{0 \leq t \leq 1} \|z(1)\| \leq C_3(R) \|w\|_{\mathcal{X}_1}^2 \leq C_4(R) \|u_0 - \hat{u}_0\|^2.$$

Since  $\|u_0 - \hat{u}_0\| \leq d$ , choosing  $d > 0$  sufficiently small ensures that

$$\|z(1)\| \leq \frac{q}{2} \|u_0 - \hat{u}_0\|.$$

Combining this with the first inequality of (3.11) in which  $v_0 = u_0 - \hat{u}_0$ , we get (3.2). This completes the proof of the theorem.

## 4 Proof of Theorem 2.1

We first outline the main steps. A well-known dissipativity argument shows that the random dynamical system defined by (2.1) has a compact invariant absorbing set  $X \subset H$ . Therefore it suffices to prove the uniqueness of an invariant measure and the property of exponential mixing for the restriction of  $(u_k, \mathbb{P}_u)$  to  $X$ . This will be done with the help of Proposition 2.4. Namely, we shall prove Proposition 2.5 and use relations (2.16) and (2.17) to define a Markov chain  $(\mathbf{u}_k, \mathbb{P}_{\mathbf{u}})$  in the extended phase space  $\mathbf{X} = X \times X$ . This Markov chain is an extension of  $(u_k, \mathbb{P}_u)$  and possesses the recurrence and exponential squeezing properties of Section 2.2, and therefore the hypotheses of Proposition 2.4 are satisfied. This will complete the proof of Theorem 2.1.

## 4.1 Reduction to a compact phase space

Well-known properties of the resolving operator for the Navier–Stokes system imply that  $S$  satisfies the inequality

$$\|S(u, f)\| \leq \varkappa \|u\| + C_1 \|f\|_{L^2(D_1)} \quad \text{for } u \in B_H(R), f \in L^2(D_1)$$

where  $\varkappa < 1$  and  $C_1 > 0$  are some universal constants. Let  $r > 0$  be so large  $\|h + \eta_k\|_{L^2(D_1)} \leq r$  almost surely. Then, with probability 1, we have

$$\|S(u, h + \eta_k)\| \leq \varkappa \rho + C_1 r \quad \text{for } u \in B_H(\rho). \quad (4.1)$$

It follows that if  $R \geq \frac{C_1 r}{1-\varkappa}$ , then the ball  $B_H(R)$  is invariant for the Markov chain  $(u_k, \mathbb{P}_u)$ . Let us take  $R = \frac{2C_1 r}{1-\varkappa}$  and denote by  $X$  the image of the set  $B_H(R) \times B_{L^2(D_1)}(r)$  under the mapping  $S$ . Then  $X$  is an invariant subset for  $(u_k, \mathbb{P}_u)$ , and the regularising property of the Navier–Stokes dynamics implies that  $X$  is compact in  $H$ . Iterating (4.1), we see that

$$\mathbb{P}_u\{u_k \in X \text{ for } k \geq k_0(\rho)\} = 1 \quad \text{for any } u \in B_H(\rho),$$

where  $k_0(\rho) = (\ln \rho + C_2)/\ln \varkappa^{-1}$  with a large constant  $C_2 > 0$ . It follows that  $(u_k, \mathbb{P}_u)$  has at least one stationary measure  $\mu$ , and any such measure is supported by  $X$ . It is easy to see that to prove (2.8), it suffices to establish inequality (2.15). The latter is proved in the next three subsections.

## 4.2 Proof of Proposition 2.5

We shall apply Proposition 5.3 to construct a measurable coupling  $(V, V')$  and Proposition 5.2 to prove (2.16). Namely, fix  $R > 0$  so large that  $X \subset B_H(R-1)$  and  $\|h + \eta_k\|_{H^1(D_1)} \leq R-1$  almost surely. Denote by  $C, d > 0$  and  $m \geq 1$  the parameters constructed in Theorem 3.1 with  $q = 1/4$ , define the Polish space

$$Z = \{(u, u') \in X \times X : \|u - u'\| \leq d\},$$

and introduce the function  $\varepsilon(u, u') = \frac{1}{2}\|u - u'\|$  on the space  $Z$ . Let us consider the pair of measures  $(P_1(u, \cdot), P_1(u', \cdot))$ . By Proposition 5.3 with  $\theta = \frac{1}{2}$ , there is a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and measurable functions  $V, V' : \Omega \times Z \rightarrow H$  such that for any  $(u, u') \in Z$  the pair  $(V(u, u'), V'(u, u'))$  is a coupling for  $(P_1(u, \cdot), P_1(u', \cdot))$  and

$$\mathbb{P}\{\|V(u, u') - V'(u, u')\| > \frac{1}{2}\|u - u'\|\} \leq C_{\|u-u'\|/4}(P_1(u, \cdot), P_1(u', \cdot)). \quad (4.2)$$

In view of (5.6), to estimate the right-hand side of this inequality, it suffices to bound the function  $K_{\|u-u'\|/4}(P_1(u, \cdot), P_1(u', \cdot))$ .

To this end, we shall apply Propositions 5.2, 5.6 and Theorem 3.1. Let us endow the ball  $B_R \subset H^1(D_1)$  with the law  $\lambda$  of the random variables  $\eta_k$ . Then  $P_1(u, \cdot)$  is the image of  $\lambda$  under the mapping  $\eta \mapsto S(u, h + \eta)$  acting from  $B_R$  to  $H$ . Let  $\mathcal{E}_m$  be the subspace entering Theorem 3.1 and let

$$\Phi : B_R \times B_H(R) \rightarrow \mathcal{L}(H, \mathcal{E}_m)$$

be a smooth  $C$ -Lipschitz mapping such that

$$\|S(u, h + \eta) - S(u', h + \eta + \Phi(h + \eta, u)(u - u'))\| \leq \frac{1}{4}\|u - u'\| \quad (4.3)$$

for any  $\eta \in B_R$  and any  $u, u' \in B_H(R)$  satisfying the inequality  $\|u - u'\| \leq d$ . The existence of such mapping was established in Theorem 3.1. Let us define a transformation  $\Psi = \Psi_{u, u'}$  of the space  $H^1(D_1)$  by the relation

$$\Psi(\eta) = \eta + \chi(\|h + \eta\|_1)\Phi(h + \eta, u)(u - u'),$$

where  $\chi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is a smooth function such that  $\chi(r) = 1$  for  $r \leq R - 1$  and  $\chi(r) = 0$  for  $r \geq R$ . The choice of the constant  $R$  and inequality (4.3) imply that

$$\|S(u, h + \eta) - S(u', h + \Psi_{u, u'}(\eta))\| \leq \frac{1}{4}\|u - u'\| \quad \text{for } \lambda\text{-a.e. } \eta \in H^1(D_1).$$

Therefore, by Proposition 5.2, we have

$$K_{\|u - u'\|/4}(P_1(u, \cdot), P_1(u', \cdot)) \leq 2\|\lambda - \Psi_*(\lambda)\|_{\text{var}}. \quad (4.4)$$

Now note that the mapping  $\Psi$  satisfies the hypotheses of Proposition 5.6 with a constant  $\varkappa$  proportional to  $\|u - u'\|$ . Combining (4.2), (4.4), and (5.22), we arrive at the required inequality (2.16).

### 4.3 Recurrence

Let us recall that the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and the  $\mathbf{X}$ -valued Markov chain  $\{\mathbf{u}_k = (u_k, u'_k)\}$  on it were defined in Section 2.2. We shall denote by  $\mathbb{P}_{\mathbf{u}}$  the probability measure associated with the initial condition  $\mathbf{u}$ . Let us set

$$\mathbf{B} = \{\mathbf{u} = (u, u') \in \mathbf{X} : \|u - u'\| \leq d\},$$

where  $d > 0$  is a small constant to be chosen later. In this and next subsections, we shall prove that the recurrence and exponential squeezing properties are satisfied for  $(\mathbf{u}_k, \mathbb{P}_{\mathbf{u}})$  with the above choice of  $\mathbf{B}$ .

To prove that  $\tau(\mathbf{B})$  is  $\mathbb{P}_{\mathbf{u}}$ -almost surely finite and satisfies (2.11), it suffices to show that

$$p := \sup_{\mathbf{u} \in \mathbf{X}} \mathbb{P}_{\mathbf{u}}\{\tau(\mathbf{B}) > \ell\} < 1, \quad (4.5)$$

where  $\ell \geq 1$  is an integer. Indeed, if this inequality is proved, then using the Markov property, for any integer  $m \geq 1$  we can write

$$\begin{aligned} \mathbb{P}_{\mathbf{u}}\{\tau(\mathbf{B}) > m\ell\} &= \mathbb{E}_{\mathbf{u}}\mathbb{P}_{\mathbf{u}}\{\tau(\mathbf{B}) > m\ell \mid \mathcal{F}_{(m-1)\ell}\} \\ &= \mathbb{E}_{\mathbf{u}}(I_{\tau(\mathbf{B}) > (m-1)\ell} \mathbb{P}_{\mathbf{u}_{(m-1)\ell}}\{\tau(\mathbf{B}) > \ell\}) \\ &\leq p \mathbb{P}_{\mathbf{u}}\{\tau(\mathbf{B}) > (m-1)\ell\}, \end{aligned}$$

where  $\{\mathcal{F}_k\}$  stands for the filtration generated by the Markov family  $(\mathbf{u}_k, \mathbb{P}_{\mathbf{u}})$ . By iteration, the above inequality implies that

$$\mathbb{P}_{\mathbf{u}}\{\tau(\mathbf{B}) > m\ell\} \leq p^m \quad \text{for all } m \geq 1. \quad (4.6)$$

A simple application of the Borel–Cantelli lemma now implies  $\tau(\mathbf{B})$  is  $\mathbb{P}_{\mathbf{u}}$ -almost surely finite. Furthermore, inequality (4.6) immediately implies that  $\tau(\mathbf{B})$  satisfies (2.11).

We now prove (4.5). Let  $\eta$  and  $\eta'$  be the random variables entering the definition of the coupling operator  $\mathcal{R}$  (see (2.17)) and let

$$\zeta_k(\omega) = \eta(\omega_k), \quad \zeta'_k(\omega) = \eta'(\omega_k), \quad k \geq 1.$$

Then  $\{\zeta_k, \zeta'_k, k \geq 1\}$  is a family of i.i.d. random variables in  $L^2(D_1)$  whose law coincides with that of  $\eta_k$ . Let  $\hat{u} \in H$  be the point defined in Hypothesis (H2). Since  $X$  is a closed absorbing subset, we have  $\hat{u} \in X$ . Using the definitions of  $\mathcal{R}$  and  $\tau$ , we can write

$$\begin{aligned} \mathbb{P}_{\mathbf{u}}\{\tau(\mathbf{B}) > \ell\} &= \mathbb{P}_{\mathbf{u}}\{\tau(\mathbf{B}) > \ell, \|u_\ell - u'_\ell\| > d\} \\ &= \mathbb{P}_{\mathbf{u}}\{\tau(\mathbf{B}) > \ell, \|S_\ell(u, \zeta_1, \dots, \zeta_\ell) - S_\ell(u, \zeta'_1, \dots, \zeta'_\ell)\| > d\} \\ &\leq \mathbb{P}\{\|S_\ell(u, \zeta_1, \dots, \zeta_\ell) - S_\ell(u', \zeta'_1, \dots, \zeta'_\ell)\| > d\} \\ &= 1 - \mathbb{P}\{\|S_\ell(u, \zeta_1, \dots, \zeta_\ell) - S_\ell(u', \zeta'_1, \dots, \zeta'_\ell)\| \leq d\} \\ &\leq 1 - P(\ell, d)^2, \end{aligned}$$

where the operator  $S_\ell$  is defined in Hypothesis (H2), and we set

$$P(\ell, d) = \inf_{v \in X} \mathbb{P}\{\|S_\ell(v, \zeta_1, \dots, \zeta_\ell) - \hat{u}\| \leq d/2\}.$$

Thus, inequality (4.5) will be established if we show that  $P(\ell, d) > 0$  for a sufficiently large integer  $\ell$ .

To this end, let us denote by  $\ell$  the integer defined in Hypothesis (H2) with  $\varepsilon = d/4$  and  $R > 0$  so large that  $X \subset B_H(R)$ . We fix  $v \in X$  and denote by  $\zeta_k = \zeta_k^v \in \mathcal{K}$  some vectors for which (2.6) holds. Let  $\delta > 0$  be so small that for any  $\tilde{\zeta}_k \in \mathcal{K}$ ,  $k = 1, \dots, \ell$ , satisfying the inequalities  $\|\tilde{\zeta}_k - \zeta_k^v\| < \delta$  we have

$$\|S_\ell(v, \tilde{\zeta}_1, \dots, \tilde{\zeta}_\ell) - \hat{u}\| \leq d/2.$$

Define the event  $\Gamma(v, \ell) = \{\|\zeta_k - \zeta_k^v\| < \delta, k = 1, \dots, \ell\} \subset \Omega$ . What has been said implies that

$$\|S_\ell(v, \zeta_1, \dots, \zeta_\ell) - \hat{u}\| \leq d/2 \quad \text{for } \omega \in \Gamma(v, \ell).$$

We see that

$$\begin{aligned} P(\ell, d) &\geq \inf_{v \in X} \mathbb{P}(\Gamma(v, \ell)) \geq \inf_{v \in X} \prod_{k=1}^{\ell} \mathbb{P}\{\|\zeta_k - \zeta_k^v\| < \delta\} \\ &\geq \left( \inf_{\xi \in \mathcal{K}} \mathbb{P}\{\|\zeta_1 - \xi\| < \delta\} \right)^\ell, \end{aligned} \quad (4.7)$$

where we used the fact that  $\zeta_k$  are i.i.d. random variables. It remains to note that the function  $\xi \mapsto \mathbb{P}\{\|\zeta_1 - \xi\| < \delta\}$  is lower semicontinuous and strictly positive on  $\mathcal{K}$ , and therefore the right-hand side of (4.7) is a positive constant. This completes the verification of the recurrence property.

#### 4.4 Exponential squeezing

Let us take arbitrary initial points  $u, u' \in X$  such that  $(u, u') \in \mathbf{B}$ . Then, by (2.16) and (2.17), we have

$$\mathbb{P}_{\mathbf{u}}\{\|u_1 - u'_1\| \leq \frac{1}{2}\|u - u'\|\} \geq 1 - C\|u - u'\|. \quad (4.8)$$

For any integer  $n \geq 1$ , let us set

$$\Gamma_n = \{\|u_k - u'_k\| \leq \frac{1}{2}\|u_{k-1} - u'_{k-1}\| \text{ for } 1 \leq k \leq n\}.$$

Note that  $\|u_k - u'_k\| \leq 2^{-k}\|u_0 - u'_0\|$  for  $1 \leq k \leq n$  on the set  $\Gamma_n$ . Combining this observation with inequality (4.8) and the Markov property, we derive

$$\begin{aligned} \mathbb{P}_{\mathbf{u}}(\Gamma_n) &= \mathbb{E}_{\mathbf{u}}(I_{\Gamma_{n-1}} \mathbb{P}_{\mathbf{u}}\{\|u_n - u'_n\| \leq \frac{1}{2}\|u_{n-1} - u'_{n-1}\| \mid \mathcal{F}_{n-1}\}) \\ &= \mathbb{E}_{\mathbf{u}}(I_{\Gamma_{n-1}} \mathbb{P}_{\mathbf{u}_{n-1}}\{\|u_1 - u'_1\| \leq \frac{1}{2}\|u_0 - u'_0\|\}) \\ &\geq (1 - 2^{1-n}C\|u - u'\|) \mathbb{P}_{\mathbf{u}}(\Gamma_{n-1}). \end{aligned}$$

Iterating this inequality, for any  $n \geq 1$  we obtain

$$\mathbb{P}_{\mathbf{u}}(\Gamma_n) \geq \prod_{k=1}^n (1 - 2^{1-k}C\|u - u'\|) \geq 1 - C_1\|u - u'\|. \quad (4.9)$$

Let us define the Markov time  $\sigma$  by relation (2.12) with  $C = d$  and  $\beta = \ln 2$ . Then, for any  $\mathbf{u} \in \mathbf{B}$ , with  $\mathbb{P}_{\mathbf{u}}$ -probability 1 we have

$$\begin{aligned} \{\sigma = \infty\} &= \{\|u_k - u'_k\| \leq 2^{-k}d \text{ for } k \geq 1\} \\ &\supset \{\|u_k - u'_k\| \leq \frac{1}{2}\|u_{k-1} - u'_{k-1}\| \text{ for } k \geq 1\} = \bigcap_{n=1}^{\infty} \Gamma_n, \end{aligned}$$

whence it follows that

$$\mathbb{P}_{\mathbf{u}}\{\sigma = \infty\} = \lim_{n \rightarrow \infty} \mathbb{P}_{\mathbf{u}}(\Gamma_n) \geq 1 - C_1\|u - u'\|.$$

Choosing  $d > 0$  sufficiently small, we arrive at (2.13).

To prove inequality (2.14), it suffices to show that

$$\mathbb{P}_{\mathbf{u}}\{\sigma = n\} \leq C_2 2^{-n} \quad \text{for } n \geq 1, \mathbf{u} \in \mathbf{B}, \quad (4.10)$$

where the positive constant  $C$  does not depend on  $\mathbf{u}$ . To this end, note that

$$\{\sigma = n\} = \{\sigma > n-1\} \cap \{\|u_n - u'_n\| > 2^{-n}d\}.$$

Using the Markov property, inequality (4.8), and the fact that  $\|u_k - u'_k\| \leq 2^{-k}d$  on the set  $\{\sigma > k\}$ , we derive

$$\begin{aligned}\mathbb{P}_u\{\sigma = n\} &= \mathbb{E}_u(I_{\{\sigma > n-1\}} \mathbb{P}_u\{\|u_n - u'_n\| > 2^{-n}d \mid \mathcal{F}_{n-1}\}) \\ &= \mathbb{E}_u(I_{\{\sigma > n-1\}} \mathbb{P}_{u_{n-1}}\{\|u_1 - u'_1\| > 2^{-n}d\}) \\ &\leq \mathbb{E}_u(I_{\{\sigma > n-1\}} \mathbb{P}_{u_{n-1}}\{\|u_1 - u'_1\| > \tfrac{1}{2}\|u_0 - u'_0\|\}) \\ &\leq 2^{1-n}Cd \mathbb{P}_u\{\sigma > n-1\}.\end{aligned}$$

We thus obtain inequality (4.10) with  $C_2 = Cd$ . The proof of Theorem 2.1 is complete.

## 5 Appendix

### 5.1 Optimal coupling

Let  $X$  be a Polish space with a metric  $d$  and let  $\mu_1, \mu_2 \in \mathcal{P}(X)$ . Recall that a pair of  $X$ -valued random variables  $(\xi_1, \xi_2)$  is called a *coupling* for  $(\mu_1, \mu_2)$  if  $\mathcal{D}(\xi_i) = \mu_i$ ,  $i = 1, 2$ . We denote by  $\Pi(\mu_1, \mu_2)$  the set of all couplings for  $(\mu_1, \mu_2)$ . Let us fix  $\varepsilon \geq 0$  and define a symmetric function  $d_\varepsilon : X \times X \rightarrow \mathbb{R}$  by the relation

$$d_\varepsilon(u_1, u_2) = \begin{cases} 1 & \text{if } d(u_1, u_2) > \varepsilon, \\ 0 & \text{if } d(u_1, u_2) \leq \varepsilon. \end{cases} \quad (5.1)$$

**Definition 5.1.** We shall say that a coupling  $(\xi_1, \xi_2) \in \Pi(\mu_1, \mu_2)$  is  $\varepsilon$ -*optimal* if it minimizes the function  $(\zeta_1, \zeta_2) \mapsto \mathbb{E} d_\varepsilon(\zeta_1, \zeta_2)$  defined on the set of all couplings for  $(\mu_1, \mu_2)$ . That is,

$$\mathbb{E} d_\varepsilon(\xi_1, \xi_2) = \inf_{(\zeta_1, \zeta_2) \in \Pi(\mu_1, \mu_2)} \mathbb{E} d_\varepsilon(\zeta_1, \zeta_2). \quad (5.2)$$

In particular, for  $\varepsilon = 0$  we obtain the usual concept of maximal coupling of measures; e.g., see [Tho00].

Let us denote by  $C_\varepsilon(\mu_1, \mu_2)$  the quantity defined by the right-hand side of (5.2) and call it the  $\varepsilon$ -*optimal cost* for the pair  $(\mu_1, \mu_2)$ . Note that, in view of (5.1) and (5.2), for any  $\varepsilon$ -optimal coupling  $(\xi_1, \xi_2)$  we have

$$\mathbb{P}\{d(\xi_1, \xi_2) > \varepsilon\} = C_\varepsilon(\mu_1, \mu_2). \quad (5.3)$$

Thus, to estimate the probability of the event that the distance between the components of an  $\varepsilon$ -optimal coupling for  $(\mu_1, \mu_2)$  is larger than  $\varepsilon$  it suffices to estimate the corresponding  $\varepsilon$ -optimal cost. We now establish a simple result that enables one to do it. Let us introduce the following function on the space  $\mathcal{P}(X) \times \mathcal{P}(X)$ :

$$K_\varepsilon(\mu_1, \mu_2) = \sup_{f, g} ((f, \mu_1) - (g, \mu_2)), \quad (5.4)$$

where the supremum is taken over all functions  $f, g \in C_b(X)$  satisfying the inequality

$$f(u_1) - g(u_2) \leq d_\varepsilon(u_1, u_2) \quad \text{for } u_1, u_2 \in X. \quad (5.5)$$

Then, by the Kantorovich duality (see Theorem 5.10 in [Vil09]), we have

$$K_\varepsilon(\mu_1, \mu_2) = C_\varepsilon(\mu_1, \mu_2) \quad (5.6)$$

Thus, to estimate the  $\varepsilon$ -optimal cost it suffices to estimate the function  $K_\varepsilon$ . The following proposition reduces this question to a “control” problem.

**Proposition 5.2.** *Let  $X$  be a compact metric space with metric  $d$ , let  $U_1, U_2$  be two  $X$ -valued random variables defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , and let  $\mu_1, \mu_2$  be their laws. Suppose there is a measurable mapping  $\Psi : \Omega \rightarrow \Omega$  such that*

$$d(U_1(\omega), U_2(\Psi(\omega))) \leq \varepsilon \quad \text{for almost every } \omega \in \Omega, \quad (5.7)$$

where  $\varepsilon > 0$  is a constant. Then

$$K_\varepsilon(\mu_1, \mu_2) \leq 2 \|\mathbb{P} - \Psi_*(\mathbb{P})\|_{\text{var}}. \quad (5.8)$$

In particular, any  $\varepsilon$ -optimal coupling  $(\xi_1, \xi_2)$  for the pair  $(\mu_1, \mu_2)$  satisfies the inequality

$$\mathbb{P}\{d(\xi_1, \xi_2) > \varepsilon\} \leq 2 \|\mathbb{P} - \Psi_*(\mathbb{P})\|_{\text{var}}. \quad (5.9)$$

*Proof.* Inequality (5.9) is a straightforward consequence of (5.8), (5.6), and (5.3). To prove (5.8), we use an argument applied in the proof of Lemma 11.8.6 in [Dud02]. Namely, note that if  $f, g \in C_b(X)$  are such that (5.5) holds, then the function  $h(u) = \sup_{v \in X} (f(v) - d_\varepsilon(u, v))$  satisfies the inequalities  $h \leq g$  and

$$f(u_1) - h(u_2) \leq d_\varepsilon(u_1, u_2), \quad |h(u_1) - h(u_2)| \leq 1 \quad \text{for } u_1, u_2 \in X.$$

It follows that

$$\begin{aligned} (f, \mu_1) - (g, \mu_2) &= \mathbb{E}(f(U_1) - g(U_2)) \\ &\leq \mathbb{E}(f(U_1) - h(U_2 \circ \Psi)) + \mathbb{E}(h(U_2 \circ \Psi) - h(U_2)) \\ &\leq \mathbb{E}d_\varepsilon(U_1, U_2 \circ \Psi) + (h \circ U_2, \Psi_*(\mathbb{P})) - (h \circ U_2, \mathbb{P}) \\ &\leq |(h \circ U_2, \mathbb{P}) - (h \circ U_2, \Psi_*(\mathbb{P}))|. \end{aligned}$$

It remains to take the supremum over  $f, g$  and to note that right-hand side of this inequality does not exceed that of (5.8).  $\square$

We now study the question of existence and measurability of an  $\varepsilon$ -optimal coupling. Let  $(Z, \mathcal{B})$  be a measurable space and let  $\{\mu_i^z, z \in Z\}$ ,  $i = 1, 2$ , be two families of probability measures on  $X$  such that the mapping  $z \mapsto \mu_i^z$  is measurable from  $Z$  to the space  $\mathcal{P}(X)$  endowed with the topology of weak convergence and the corresponding  $\sigma$ -algebra. The following proposition establishes the existence of an “almost”  $\varepsilon$ -optimal coupling that is a measurable function of  $z$ , provided that  $X$  is a subset of a Banach space.

**Proposition 5.3.** *In addition to the above hypotheses, assume that  $X$  is a compact subset of a Banach space  $Y$ . Then for any positive measurable function  $\varepsilon(z)$ ,  $z \in Z$ , and any  $\theta \in (0, 1)$  there is a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and measurable functions  $\xi_i^z(\omega) : \Omega \times Z \rightarrow Y$ ,  $i = 1, 2$ , such that, for any  $z \in Z$ , the pair  $(\xi_1^z, \xi_2^z)$  is a coupling for  $(\mu_1^z, \mu_2^z)$  and*

$$\mathbb{E} d_{\varepsilon(z)}(\xi_1^z, \xi_2^z) \leq C_{\theta \varepsilon(z)}(\mu_1^z, \mu_2^z). \quad (5.10)$$

*Proof.* The existence of a measurable optimal  $\varepsilon(z)$ -coupling would be an immediate consequence of Corollary 5.22 in [Vil09] if the function  $d_{\varepsilon(z)}(u_1, u_2)$  was continuous and independent of  $z$ . Since this is not the case, we now outline the modifications that are needed to prove the existence of a measurable coupling satisfying (5.10).

Let us write  $Z$  as the union of countably many disjoint measurable subset  $Z_k$  such that the image of the restriction of  $\varepsilon(z)$  to any of them is contained in a bounded interval separated from zero. For instance, we can take  $Z_k = \{z \in Z : (k+1)^{-1} < \varepsilon(z) \leq k^{-1}\}$  with  $k \geq 1$  and use a similar partition for  $(1, +\infty)$ . If we construct measurable couplings on each set  $Z_k$  for which (5.10) holds with  $Z = Z_k$ , then by gluing them together, we get the required random variables  $\xi_i^z$ ,  $i = 1, 2$ .

To construct a measurable coupling on  $Z_k$ , we first reduce the problem to the case when  $\varepsilon \equiv 1$ . Let us consider stretched measures  $\tilde{\mu}_i^z$  defined by the formula  $\tilde{\mu}_i^z(\Gamma) = \mu_i^z(\varepsilon(z)\Gamma)$  for  $\Gamma \in \mathcal{B}_Y$ . The measures  $\tilde{\mu}_i^z$  are supported by a compact set  $X_k \subset Y$ , and if  $(\tilde{\xi}_1^z, \tilde{\xi}_2^z)$  is a measurable coupling such that

$$\mathbb{E} d_1(\tilde{\xi}_1^z, \tilde{\xi}_2^z) \leq C_{\theta}(\tilde{\mu}_1^z, \tilde{\mu}_2^z) \quad \text{for any } z \in Z_k,$$

then  $\varepsilon(z)(\tilde{\xi}_1^z, \tilde{\xi}_2^z)$  is a measurable coupling for the original measures that satisfies (5.10). Thus, we can assume from the very beginning that  $\varepsilon \equiv 1$ .

Let  $d : X \times X \rightarrow \mathbb{R}$  be an arbitrary continuous symmetric function such that

$$d_1(u_1, u_2) \leq d(u_1, u_2) \leq d_{\theta}(u_1, u_2) \quad \text{for all } u_1, u_2 \in X. \quad (5.11)$$

By Corollary 5.22 in [Vil09], there is a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and measurable functions  $\xi_i^z : Z \times \Omega \rightarrow X$  such that  $(\xi_1^z, \xi_2^z)$  is a coupling for  $(\mu_1^z, \mu_2^z)$  for any  $z \in Z$ , and

$$\mathbb{E} d(\xi_1^z, \xi_2^z) = \inf_{(\zeta_1, \zeta_2) \in \Pi(\mu_1^z, \mu_2^z)} \mathbb{E} d(\zeta_1, \zeta_2).$$

Combining this relation with (5.11), we arrive at (5.10) with  $\varepsilon \equiv 1$ .  $\square$

## 5.2 Truncated observability inequality

Let  $\delta \in (0, 1)$  and let  $\hat{u} \in \mathcal{Y}_{\delta, 1}$  be an arbitrary function. Consider the adjoint equation for the Navier–Stokes system linearised around  $\hat{u}$ :

$$\dot{g} + \langle \hat{u}, \nabla \rangle g + \langle \nabla g, \hat{u} \rangle + \nu \Delta g + \nabla \pi = 0, \quad \operatorname{div} g = 0. \quad (5.12)$$

Let us fix an open set  $Q \subset (\delta, 1) \times D$ , an orthonormal basis  $\{\varphi_j\} \subset L^2(Q, \mathbb{R}^2)$ , and a function  $\chi \in C_0^\infty(Q)$  and denote by  $P_m$  the orthogonal projection in

$L^2(Q, \mathbb{R}^2)$  onto the  $m$ -dimensional subspace spanned by  $\varphi_j$ ,  $j = 1, \dots, m$ . Recall that  $H_N$  stands for the vector span of the first  $N$  eigenfunctions of the Stokes operator. The following result is a simple consequence of the observability inequality for the linearised Navier–Stokes system (see [FE99, Ima01, FCGIP04]).

**Proposition 5.4.** *For any  $\delta, \rho > 0$  and any integer  $N \geq 1$  there is an integer  $m \geq 1$  such that if  $\hat{u} \in B_{\mathcal{Y}_{\delta,1}}(\rho)$ , then any solution  $g \in \mathcal{X}_1$  of Eq. (5.12) with  $g(1) \in H_N$  satisfies the inequality*

$$\|g(0)\| \leq C \|\mathbf{P}_m(\chi g)\|_{L^2(D_1)}, \quad (5.13)$$

where  $C > 0$  is a constant depending only on  $\delta$  and  $\rho$ .

*Proof.* We essentially repeat the argument used in the paper [BRS11] in which inequality (5.13) is proved for a time-independent function  $\chi$  and a projection  $\mathbf{P}_m$  acting in the space variables. We claim that if  $g \in \mathcal{X}_1$  is a solution of (5.12) with  $g(1) \in H_N$ , then

$$\|\chi g\|_{H^1(D_1)} \leq C_1 \|\chi g\|_{L^2(D_1)}, \quad (5.14)$$

where  $C_1 > 0$  is a constant depending on  $\delta$ ,  $\rho$ , and  $N$ . Once this inequality is established, the required result can be derived by using the fact that

$$\|(I - \mathbf{P}_m)v\| \leq \delta_m \|v\|_1 \quad \text{for any } v \in H_0^1(Q, \mathbb{R}^2),$$

where  $\{\delta_m\}$  is a sequence going to zero as  $m \rightarrow \infty$ ; cf. proof of Proposition 5.3 in [BRS11].

Suppose that (5.14) is false. Then there are functions  $\hat{u}_n \in B_{\mathcal{Y}_{\delta,1}}(\rho)$  and solutions  $g_n \in \mathcal{X}_1$  of (5.12) with  $\hat{u} = \hat{u}_n$  such that

$$g_n(1) \in H_N, \quad \|g_n(1)\| = 1, \quad (5.15)$$

$$\|\chi g_n\|_{H^1(D_1)} \geq n \|\chi g_n\|_{L^2(D_1)}. \quad (5.16)$$

Now note that  $\|g_n(1)\|_{H^2} \leq C_2$  for all  $n \geq 1$ , whence it follows, by standard estimates for the 2D Navier–Stokes system, that

$$\|g_n\|_{\mathcal{X}_1} + \|g_n\|_{L^2(J, H^3)} + \|\partial_t g_n\|_{L^2(J, V)} \leq C_3, \quad n \geq 1.$$

Passing to a subsequence, we can assume that  $\{\hat{u}_n\}$  and  $\{g_n\}$  converge weakly (in appropriate functional spaces) to some functions  $\hat{u}$  and  $g$ , respectively, such that  $\hat{u} \in B_{\mathcal{Y}_{\delta,1}}(\rho)$ ,  $g \in \mathcal{X}_1$  is a solution of (5.12), and  $\|g(1)\| = 1$ . Moreover, it follows from (5.16) that  $\chi g \equiv 0$ . Let us show that the latter is impossible.

Indeed, let an interval  $(a, b) \subset J$  and a ball  $B \subset D$  be such that

$$\chi(t, x) \geq \alpha > 0 \quad \text{for } (t, x) \in (a, b) \times B.$$

Then, by the observability inequality (see Lemma 1 in [FCGIP04]), we obtain

$$\|g(a)\| \leq C_4 \|g\|_{L^2((a,b) \times D)} \leq C_4 \alpha^{-1} \|\chi g\|_{L^2(D_1)} = 0.$$

The backward uniqueness for the linearised Navier–Stokes system (see Section II.8 in [BV92]) implies that  $g(t) = 0$  for  $a \leq t \leq 1$ . This contradicts the fact that  $\|g(1)\| = 1$ . The proof is complete.  $\square$

### 5.3 Minima of second-order polynomials on Hilbert spaces

Let  $U$  and  $Y$  be real separable Hilbert spaces and let  $F : U \times Y \rightarrow \mathbb{R}$  be a function of the form

$$F(u, y) = (Q_y u, u)_U + (a_y, u)_U + b_y,$$

where  $Q_y \in \mathcal{L}(U)$  is a self-adjoint operator,  $a_y, b_y \in U$  for all  $y \in Y$ , and  $(\cdot, \cdot)_U$  stands for the scalar product in  $U$ . We assume that

$$(Q_y u, u)_U \geq c \|u\|_U^2 \quad \text{for all } u \in U, y \in Y. \quad (5.17)$$

In this case, it is easy to see that for any  $y \in Y$  the function  $u \mapsto F(u, y)$  has a unique global minimum  $u^* = u^*(y)$ . The following simple proposition establishes some properties of  $u^*$ .

**Proposition 5.5.** *Let us assume that the functions*

$$y \mapsto Q_y, \quad y \mapsto (a_y, b_y) \quad (5.18)$$

*are infinitely smooth from  $Y$  to the spaces  $\mathcal{L}(U)$  and  $U \times U$ , respectively. Then the unique minimum  $u^*(y)$  is an infinitely smooth function of  $y \in Y$ . Moreover, this implication remains true if the property of infinite smoothness is replaced by Lipschitz continuity on bounded balls.*

*Proof.* Inequality (5.17) implies that  $u^*(y)$  is the only solution of the linear equation

$$(\partial_u F)(u, y) = 0. \quad (5.19)$$

Therefore, the required smoothness of  $u^*$  will be established if we show that the implicit function theorem can be applied to (5.19). This is a straightforward consequence of the fact that  $\partial_u F$  is an infinitely smooth function of its arguments, and its derivative in  $u$  coincides with  $2Q_y$ .

We now prove that if  $a_y$ ,  $b_y$ , and  $Q_y$  are Lipschitz continuous on bounded balls, then so is  $u^*(y)$ . Indeed, fix  $R > 0$  and take two points  $y_1, y_2 \in B_Y(R)$ . It follows from (5.19) and the explicit form of the derivative  $\partial_u F$  that

$$Q_{y_1}(u^*(y_1) - u^*(y_2)) = (Q_{y_2} - Q_{y_1})u^*(y_2) + \frac{1}{2}(a_{y_2} - a_{y_1}).$$

Taking the scalar product of this equation with  $u^*(y_1) - u^*(y_2)$  and using inequality (5.17) and the Lipschitz continuity  $Q_y$  and  $a_y$ , we obtain the required result.  $\square$

### 5.4 Image of measures under finite-dimensional transformations

Let  $X$  be a separable Banach space endowed with a norm  $\|\cdot\|$  and represented as the direct sum

$$X = E \dot{+} F, \quad (5.20)$$

where  $E \subset X$  is a finite-dimensional subspace. Denote by  $\mathbf{P}_E$  and  $\mathbf{P}_F$  the projections corresponding to decomposition (5.20). Let  $\lambda \in \mathcal{P}(X)$  be a measure that can be written as the tensor product of its projections  $\lambda_E = (\mathbf{P}_E)_*\lambda$  and  $\lambda_F = (\mathbf{P}_F)_*\lambda$ . Assume also that  $\lambda$  has a bounded support and that  $\lambda_E$  possesses a  $C^1$ -smooth density  $\rho$  with respect to the Lebesgue measure on  $E$ . The following simple result gives an estimate for the total variation distance between the measure  $\lambda$  and its image under a diffeomorphism of  $X$  acting only along  $E$ .

**Proposition 5.6.** *Under the above hypotheses, let  $\Psi : X \rightarrow X$  be a transformation that can be written in the form*

$$\Psi(u) = u + \Phi(u), \quad u \in X,$$

where  $\Phi$  is a  $C^1$ -smooth map such that the image of  $\Phi$  is contained in  $E$  and

$$\|\Phi(u_1)\| \leq \varkappa, \quad \|\Phi(u_1) - \Phi(u_2)\| \leq \varkappa \|u_1 - u_2\| \quad \text{for all } u_1, u_2 \in X. \quad (5.21)$$

Then the total variation distance between  $\lambda$  and its image under  $\Psi$  admits the estimate

$$\|\lambda - \Psi_*(\lambda)\|_{\text{var}} \leq C\varkappa, \quad (5.22)$$

where  $C > 0$  is a constant not depending on  $\varkappa$ .

This proposition is a simple particular case of more general results presented in Chapter 10 of [Bog10]. However, for the reader's convenience, we give a complete proof of Proposition 5.6.

*Proof.* Inequality (5.22) needs to be proved for  $\varkappa \ll 1$ , because it is trivial if  $\varkappa$  is separated from zero. Furthermore, since  $\lambda$  has a bounded support, there is no loss of generality in assuming that  $\Phi$  vanishes outside a large ball.

We first note that if  $f \in C_b(X)$ , then

$$(f, \lambda) = \int_F \lambda_F(dw) \int_E f(v+w) \rho(v) dv.$$

Since  $v \mapsto v + \Phi(v+w)$  is a  $C^1$ -diffeomorphism of  $E$  for  $|\varkappa| \ll 1$ , we see that

$$\begin{aligned} (f, \Psi_*(\lambda)) &= \int_F \lambda_F(dw) \int_E f(v+w+\Phi(v+w)) \rho(v) dv \\ &= \int_F \lambda_F(dw) \int_E \frac{f(v'+w) \rho(\Theta_w(v'))}{\det(I + (D\Phi)(\Theta_w(v') + w))} dv', \end{aligned}$$

where  $\Theta_w(v')$  denotes the solution of the equation  $v + \Phi(v+w) = v'$  and  $D\Phi$  stands for the Fréchet derivative of  $\Phi$ . Combining the above formulas, we get

$$\delta(f, \lambda) := (f, \Psi_*(\lambda)) - (f, \lambda) = \int_F \lambda_F(dw) \int_E f(v+w) \Delta(v, w) dv, \quad (5.23)$$

where we set

$$\Delta(v, w) = \frac{\rho(\Theta_w(v))}{\det(I + (D\Phi)(\Theta_w(v) + w))} - \rho(v).$$

The  $\varkappa$ -Lipschitz continuity of  $\Phi$  implies that

$$|\det(I + (D\Phi)(\Theta_w(v) + w))^{-1} - 1| = O(\varkappa) \quad \text{for all } v \in E, w \in F,$$

where  $O(\varkappa)$  denotes any function whose absolute value can be estimated by  $C\varkappa$  (uniformly with respect to all other variables). It follows that

$$\Delta(v, w) = \frac{\rho(\Theta_w(v)) - \rho(v)}{\det(I + (D\Phi)(\Theta_w(v) + w))} + \rho(v)O(\varkappa). \quad (5.24)$$

Furthermore, using inequalities (5.21), we derive

$$\|\Theta_w(v)\| \leq \frac{1}{1-\varkappa} (\|v\| + \varkappa\|w\|), \quad \|\Theta_w(v) - v\| \leq \frac{\varkappa}{1-\varkappa} (\|v\| + \|w\|). \quad (5.25)$$

Substituting (5.24) into (5.23) and using inequalities (5.25), we obtain

$$\begin{aligned} |\delta(f, \lambda)| &\leq \int_F \lambda_F(dw) \int_E |f(v+w)| \frac{|\rho(\Theta_w(v)) - \rho(v)|}{\det(I + (D\Phi)(\Theta_w(v) + w))} dv + O(\varkappa) \\ &\leq \|f\|_\infty \int_F \lambda_F(dw) \int_E |\rho(v) - \rho(v + \Phi(v+w))| dv + O(\varkappa). \end{aligned}$$

Taking the supremum over  $f \in C_b(X)$  satisfying  $\|f\|_\infty \leq 1$  and noting that the integrals on the right-hand side can be taken over sufficiently large balls, we obtain

$$\begin{aligned} \|\lambda - \Psi_*(\lambda)\|_{\text{var}} &\leq \frac{1}{2} \int_{B_F(R)} \lambda_F(dw) \int_{B_F(R)} |\rho(v) - \rho(v + \Phi(v+w))| dv + O(\varkappa) \\ &\leq \frac{\varkappa}{2} \int_{B_F(R)} \lambda_F(dw) \int_{B_F(R)} \int_0^1 |\nabla \rho(v + \theta\Phi(v+w))| d\theta dv + O(\varkappa), \end{aligned}$$

where we used the first inequality in (5.21). It is straightforward to see that the right-hand side of the above inequality can be estimated by  $C\varkappa$ .  $\square$

**Acknowledgment.** The author is grateful to Lihu Xu for stimulating discussion.

## References

- [Bar03] V. Barbu, *Feedback stabilization of Navier–Stokes equations*, ESAIM Control Optim. Calc. Var. **9** (2003), 197–206 (electronic).
- [BKL02] J. Bricmont, A. Kupiainen, and R. Lefevre, *Exponential mixing for the 2D stochastic Navier–Stokes dynamics*, Comm. Math. Phys. **230** (2002), no. 1, 87–132.

- [Bog10] V. I. Bogachev, *Differentiable Measures and the Malliavin Calculus*, American Mathematical Society, Providence, RI, 2010.
- [BRS11] V. Barbu, S. S. Rodrigues, and A. Shirikyan, *Internal exponential stabilization to a non-stationary solution for 3D Navier–Stokes equations*, SIAM J. Control Optim. **49** (2011), no. 4, 1454–1478.
- [BT04] V. Barbu and R. Triggiani, *Internal stabilization of Navier–Stokes equations with finite-dimensional controllers*, Indiana Univ. Math. J. **53** (2004), no. 5, 1443–1494.
- [BV92] A. Babin and M. Vishik, *Attractors of Evolutionary Equations*, North-Holland, Amsterdam, 1992.
- [Dud02] R. M. Dudley, *Real Analysis and Probability*, Cambridge University Press, Cambridge, 2002.
- [EMS01] W. E, J. C. Mattingly, and Ya. G. Sinai, *Gibbsian dynamics and ergodicity for the stochastically forced Navier–Stokes equation*, Comm. Math. Phys. **224** (2001), 83–106.
- [FCGIP04] E. Fernández-Cara, S. Guerrero, O. Yu. Imanuvilov, and J.-P. Puel, *Local exact controllability of the Navier–Stokes system*, J. Math. Pures Appl. (9) **83** (2004), no. 12, 1501–1542.
- [FE99] A. V. Fursikov and O. Yu. Emanuilov, *Exact controllability of the Navier–Stokes and Boussinesq equations*, Uspekhi Mat. Nauk **54** (1999), no. 3 (327), 93–146.
- [FM95] F. Flandoli and B. Maslowski, *Ergodicity of the 2D Navier–Stokes equation under random perturbations*, Comm. Math. Phys. **172** (1995), 119–141.
- [Fur04] A. V. Fursikov, *Stabilization for the 3D Navier–Stokes system by feedback boundary control*, Discrete Contin. Dyn. Syst. **10** (2004), no. 1-2, 289–314.
- [Hai02] M. Hairer, *Exponential mixing properties of stochastic PDE’s through asymptotic coupling*, Probab. Theory Related Fields **124** (2002), 345–380.
- [HM06] M. Hairer and J. Mattingly, *Ergodicity of the 2D Navier–Stokes equations with degenerate stochastic forcing*, Ann. of Math. (2) **164** (2006), no. 3, 993–1032.
- [HM08] ———, *Spectral gaps in Wasserstein distances and the 2D stochastic Navier–Stokes equations*, Ann. Probab. **36** (2008), no. 6, 2050–2091.
- [Hör63] L. Hörmander, *Linear Partial Differential Operators*, Academic Press, New York, 1963.

- [Ima01] O. Yu. Imanuvilov, *Remarks on exact controllability for the Navier-Stokes equations*, ESAIM Control Optim. Calc. Var. **6** (2001), 39–72 (electronic).
- [IT79] A. D. Ioffe and V. M. Tihomirov, *Theory of Extremal Problems*, North-Holland, Amsterdam, 1979.
- [KPS02] S. Kuksin, A. Piatnitski, and A. Shirikyan, *A coupling approach to randomly forced nonlinear PDEs. II*, Comm. Math. Phys. **230** (2002), no. 1, 81–85.
- [KS00] S. Kuksin and A. Shirikyan, *Stochastic dissipative PDE's and Gibbs measures*, Comm. Math. Phys. **213** (2000), no. 2, 291–330.
- [KS01] ———, *A coupling approach to randomly forced nonlinear PDE's. I*, Comm. Math. Phys. **221** (2001), 351–366.
- [KS11] ———, *Mathematics of 2D Statistical Hydrodynamics*, manuscript of a book, 2011.
- [Kuk82] S. Kuksin, *Diffeomorphisms of function spaces that correspond to quasilinear parabolic equations*, Mat. Sb. (N.S.) **117(159)** (1982), no. 3, 359–378, 431.
- [Mat02] J. Mattingly, *Exponential convergence for the stochastically forced Navier–Stokes equations and other partially dissipative dynamics*, Comm. Math. Phys. **230** (2002), 421–462.
- [MY02] N. Masmoudi and L.-S. Young, *Ergodic theory of infinite dimensional systems with applications to dissipative parabolic PDE's*, Comm. Math. Phys. (2002), 461–481.
- [Shi04] A. Shirikyan, *Exponential mixing for 2D Navier–Stokes equations perturbed by an unbounded noise*, J. Math. Fluid Mech. **6** (2004), no. 2, 169–193.
- [Shi08] ———, *Exponential mixing for randomly forced partial differential equations: method of coupling*, Instability in Models Connected with Fluid Flows. II, vol. 7, Springer, New York, 2008, pp. 155–188.
- [Tem79] R. Temam, *Navier–Stokes Equations*, North-Holland, Amsterdam, 1979.
- [Tho00] H. Thorisson, *Coupling, Stationarity, and Regeneration*, Springer-Verlag, New York, 2000.
- [VF88] M. I. Vishik and A. V. Fursikov, *Mathematical Problems in Statistical Hydromechanics*, Kluwer, Dordrecht, 1988.
- [Vil09] C. Villani, *Optimal Transport. Old and New*, Springer-Verlag, Berlin, 2009.